

PF 12-APR-2001; 2001WO-IB00619.
XX
PR 13-APR-2000; 2000US-196909P.
XX
PR 01-FEB-2001; 2001US-265310P.
XX
PA (ONEO-) 1149336 ONTARIO INC.
XX
PI Drucker DJ, Lovshin JAL;
XX
WPI: 2002-026012/03.
XX
PT Novel recombinant DNA construct, useful as drug screening tool and as
PT protein delivery vehicle, comprising glucagon-like peptide-2 receptor
PT gene, and linked to heterologous gene of interest for expression
XX
PS Disclosure; Fig 1; 70pp; English.
XX

CC This sequence represents the 5' flanking and 5' UTR of the murine
CC glucagon-like peptide-2 (GLP-2) receptor gene. This sequence may be
CC used in the DNA construct of the invention, such that it is linked
CC for expression with a heterologous gene of interest. This construct
CC is useful for screening compounds to identify regulators of GLP-2
CC receptor expression. The DNA construct is also useful for delivering
CC a gene product to tissues selected from intestinal epithelium and the
CC lateral hypothalamus, by transfecting an organism, or a gametic or
CC embryonic form of such an organism, with the construct. The DNA
CC construct is also useful to identify cells capable of mediating
CC expression from the chosen GLP-2R promoter, to identify regions of
CC the GLP-2R promoter that are functional in a given cell type, and
CC to screen for agents that modulate expression from the GLP-2R promoter.
CC It can be used to map functional regions of the GLP-2R promoter.
CC for screening agents that modulate the function of GLP-2R promoter.
XX
SQ Sequence 2170 BP; 574 A; 501 C; 550 G; 540 T; 5 other;

Query Match 29.6%; Score 62.2; DB 24; Length 2170;
Best Local Similarity 67.5%; Pred. No. 1.3e-09;
Matches 141; Conservative 0; Mismatches 48; Indels 20; Gaps 3;
QY 3 GCCTTGTCTTCTCTCCTCAGCTGTCAGAGAGTCCAGAGACACAGTACTT-AGGG 61
DB 1476 GTCTTCTCTTCTCTCTGCTGAGAGAGTCCAGGACGATGAGCTTTGGGG 1535
QY 62 AAGCTTGGGAAAAATCTCCCTGCTTTTGGGGGGGCGAGGGGATGACCCAGGCC 121
DB 1536 TAGCTTGGGAAAAATCTCCAGAGATTTAGAGGGGCGAGCGGGGATGA----- 1586
QY 122 GAGAGAGAACTCTGAGACTCCGTAGATTGCTAGACCCCTCAGACTCTCGGGCA 181
DB 1587 -----GAAACTTGGAGATTGCTAGATCGCTTAGAGCACTCAACAGTCT-----GGCG 1636
QY 182 CGCTGAGAGGATTTTGCACAACTTCC 210
DB 1637 GCCTGAAGAGGACTTGTGCAACACTTCC 1665

RESULT 2

AAH04015/C
ID AAH04015 standard; CDNA; 772 BP.

AAH04015;

26-JUN-2001 (first entry)

Human cDNA clone (5'-primer) SEQ ID NO:850.

Human; primer; detection; diagnosis; antisense therapy; gene therapy; ss.

Homo sapiens.

EP1074617-A2.

XX 07-FEB-2001.

XX
PF 28-JUL-2000; 2000EP-0116126.
XX
XX 29-JUL-1999; 99JP-0248036.
XX 27-AUG-1999; 99JP-0300253.
PR 11-JAN-2000; 2000JP-0118776.
PR 02-MAY-2000; 2000JP-0183767.
PR 09-JUN-2000; 2000JP-0241899.
XX
XX (HELI-) HELIX RES INST.
XX
PI Ota T, Isogai T, Nishikawa T, Hayashi K, Saito K, Yamamoto J;
PI Ishii S, Sugiyama T, Wakamatsu A, Nagai K, Otsuki T;
XX
WPI: 2001-318749/34.
XX

PT Primer sets for synthesizing polynucleotides, particularly the 5602
PT full-length cDNAs defined in the specification, and for the detection
PT and/or diagnosis of the abnormality of the proteins encoded by the
PT full-length cDNAs.

Claim 1; SEQ ID 850; 2537bp + CD ROM; English.

CC The present invention describes primer sets for synthesizing 5602
CC full-length cDNAs defined in the specification. Where a primer set
CC comprises: (a) an oligo-dr primer and an oligonucleotide complementary
CC to the complementary strand of a polynucleotide which comprises one of
CC the 5602 nucleotide sequences defined in the specification, where the
CC oligonucleotide comprises at least 15 nucleotides; or (b) a combination
CC of an oligonucleotide comprising a sequence complementary to the
CC complementary strand of a polynucleotide which comprises a 5'-end
CC sequence and an oligonucleotide comprising a sequence complementary to a
CC polynucleotide which comprises a 3'-end sequence, where the
CC oligonucleotide comprises at least 15 nucleotides and the combination of
CC the 5'-end sequence/3'-end sequence is selected from those defined in
CC the specification. The primer sets can be used in antisense therapy and
CC in gene therapy. The primers are useful for synthesizing polynucleotides,
CC particularly full-length cDNAs. The primers are also useful for the
CC detection and/or diagnosis of the abnormality of the proteins encoded by
CC the full-length cDNAs. The primers allow obtaining of the full-length
CC cDNAs easily without any specialised methods. AAH03166 to AAH13628 and
CC AAH13633 to AAH18742 represent human cDNA sequences; AAB92446 to
CC AAB95893 represent human amino acid sequences; and AAH13629 to AAH13632
CC represent oligonucleotides, all of which are used in the exemplification
CC of the present invention.
XX

SQ Sequence 772 BP; 147 A; 258 C; 194 G; 170 T; 3 other;

Query Match 18.4%; Score 38.6; DB 22; Length 772;
Best Local Similarity 52.1%; Pred. No. 0.023;
Matches 86; Conservative 0; Mismatches 79; Indels 0; Gaps 0;

QY 37 CCCAGAAAGACACAGCTGACTTAGGGAAGTCTGGAAAAATCTCCCTTTGGGGGGG 96
DB 258 CCCAGAGAGCAGCAGACACACCTGGGAAAGACAGGAGCACTCTCCAGCCTGGCACCCCT 199

QY 97 CAGGGGCGGGGATGAGCCAGGCGCGAAGAACTCTGAAGACTCGTAGATTGCTCTA 156
DB 198 TAAATCTGAGAGAGGAAGCTCTTGGCTGGGGGGAACATTACAGAGCCTGGAGCATAGCTTA 139

QY 157 GACCGCTCAGACACTCTCTGGCGCAGCGTGGAGAGAGATTGTGCA 201
DB 138 CACCTCATAGGGAAGACAGGCGCTGATCTGCTGAGAGGTTGGCCA 94

RESULT 3

AAH15417/C

ID AAH15417 standard; CDNA; 1612 BP.

AAH15417;

26-JUN-2001 (first entry)

XX

PR 12-SEP-2000; 2000US-0231968.
 PR 14-SEP-2000; 2000US-0233397.
 PR 14-SEP-2000; 2000US-0233398.
 PR 14-SEP-2000; 2000US-0233399.
 PR 14-SEP-2000; 2000US-0232400.
 PR 14-SEP-2000; 2000US-0232401.
 PR 14-SEP-2000; 2000US-0233063.
 PR 14-SEP-2000; 2000US-0233064.
 PR 21-SEP-2000; 2000US-0233065.
 PR 21-SEP-2000; 2000US-0233066.
 PR 21-SEP-2000; 2000US-0234223.
 PR 25-SEP-2000; 2000US-0234274.
 PR 25-SEP-2000; 2000US-0234997.
 PR 26-SEP-2000; 2000US-0234998.
 PR 27-SEP-2000; 2000US-0235484.
 PR 27-SEP-2000; 2000US-0235834.
 PR 29-SEP-2000; 2000US-0235836.
 PR 29-SEP-2000; 2000US-0236327.
 PR 29-SEP-2000; 2000US-0236367.
 PR 29-SEP-2000; 2000US-0236368.
 PR 29-SEP-2000; 2000US-0236369.
 PR 02-OCT-2000; 2000US-0236370.
 PR 02-OCT-2000; 2000US-0237037.
 PR 02-OCT-2000; 2000US-0237038.
 PR 02-OCT-2000; 2000US-0237039.
 PR 13-OCT-2000; 2000US-0237040.
 PR 13-OCT-2000; 2000US-0239935.
 PR 20-OCT-2000; 2000US-0239937.
 PR 20-OCT-2000; 2000US-0240960.
 PR 20-OCT-2000; 2000US-0241221.
 PR 20-OCT-2000; 2000US-0241785.
 PR 20-OCT-2000; 2000US-0241786.
 PR 20-OCT-2000; 2000US-0241787.
 PR 20-OCT-2000; 2000US-0241808.
 PR 20-OCT-2000; 2000US-0241809.
 PR 01-NOV-2000; 2000US-0241826.
 PR 08-NOV-2000; 2000US-0244617.
 PR 08-NOV-2000; 2000US-0246474.
 PR 08-NOV-2000; 2000US-0246475.
 PR 08-NOV-2000; 2000US-0246476.
 PR 08-NOV-2000; 2000US-0246477.
 PR 08-NOV-2000; 2000US-0246478.
 PR 08-NOV-2000; 2000US-0246523.
 PR 08-NOV-2000; 2000US-0246524.
 PR 08-NOV-2000; 2000US-0246525.
 PR 08-NOV-2000; 2000US-0246526.
 PR 08-NOV-2000; 2000US-0246527.
 PR 08-NOV-2000; 2000US-0246528.
 PR 08-NOV-2000; 2000US-0246532.
 PR 08-NOV-2000; 2000US-0246609.
 PR 08-NOV-2000; 2000US-0246610.
 PR 08-NOV-2000; 2000US-0246611.
 PR 08-NOV-2000; 2000US-0246613.
 PR 17-NOV-2000; 2000US-0249207.
 PR 17-NOV-2000; 2000US-0249208.
 PR 17-NOV-2000; 2000US-0249209.
 PR 17-NOV-2000; 2000US-0249210.
 PR 17-NOV-2000; 2000US-0249211.
 PR 17-NOV-2000; 2000US-0249212.
 PR 17-NOV-2000; 2000US-0249213.
 PR 17-NOV-2000; 2000US-0249214.
 PR 17-NOV-2000; 2000US-0249215.
 PR 17-NOV-2000; 2000US-0249216.
 PR 17-NOV-2000; 2000US-0249217.
 PR 17-NOV-2000; 2000US-0249218.
 PR 17-NOV-2000; 2000US-0249244.
 PR 17-NOV-2000; 2000US-0249245.
 PR 17-NOV-2000; 2000US-0249264.
 PR 17-NOV-2000; 2000US-0249265.
 PR 17-NOV-2000; 2000US-0249297.
 PR 17-NOV-2000; 2000US-0249299.
 PR 17-NOV-2000; 2000US-0249300.
 PR 01-DEC-2000; 2000US-0250160.

PR 01-DEC-2000; 2000US-0250391.
 PR 05-DEC-2000; 2000US-0251030.
 PR 05-DEC-2000; 2000US-0251988.
 PR 06-DEC-2000; 2000US-0256719.
 PR 08-DEC-2000; 2000US-0251479.
 PR 08-DEC-2000; 2000US-0251856.
 PR 08-DEC-2000; 2000US-0251868.
 PR 08-DEC-2000; 2000US-0251869.
 PR 08-DEC-2000; 2000US-0251989.
 PR 11-DEC-2000; 2000US-0251990.
 PR 05-JAN-2001; 2001US-0254097.
 PR 05-JAN-2001; 2001US-0259678.
 PR XX
 PA (HUMA-) HUMAN GENOME SCI INC.
 PR XX
 PI Rosen CA, Barash SC, Ruben SM;
 PR XX
 DR WPI; 2001-465570/50.
 PR XX
 PT Isolated nucleic acid molecule encoding a reproductive system antigen
 PR PT is used in preventing, treating or ameliorating a medical condition -
 PR XX
 PS Disclosure; SEQ ID NO 5627; 1297bp + Sequence Listing; English.
 PR XX
 CC The present invention provides the protein and coding sequences of a
 CC number of human reproductive system related antigens. These can be used
 CC in the prevention and treatment of reproductive system disorders,
 CC including cancer. The present sequence is a genomic sequence encoding a
 CC protein of the invention.
 PR XX
 SQ Sequence 13327 BP; 3096 A; 3875 C; 3799 G; 2557 T; 0 other;
 PR XX
 QY Query Match 16.3%; Score 34.2; DB 22; Length 13327;
 PR XX
 DB Best Local Similarity 58.3%; Pred. No. 1.3; Matches 60; Conservative 0; Mismatches 43; Indels 0; Gaps 0;
 PR XX
 QY 16 TCCTGAGCCTGTCAAGAGTCCAGAAAGCAACGCTGACTTAGGAGAGCTGGGAA 75
 PR XX
 DB 5824 TCATCAGAGGACAAAGAGCCAGGAAACCTGGGAGATGGGAGAGAGACCGAGAA 75
 PR XX
 QY 76 ATCTCCCTGCTTTGGGGGGGAGGGGGGATGAGCCAGG 118
 PR XX
 DB 5884 ATGGGGCTCTCTGTGGGTACGAGAGAGTGGCTTAAGGAAG 5926
 PR XX
 RESULT 5
 PR XX
 ID AAL07185/C
 PR XX
 XX AAL07185; standard; DNA; 496 BP.
 PR XX
 AC AAL07185;
 PR XX
 XX
 PR DT 21-NOV-2001 (first entry)
 PR XX
 XX Human reproductive system related antigen DNA SEQ ID NO: 9873.
 PR DE Human reproductive system related antigen DNA SEQ ID NO: 9873.
 PR XX
 KW Human; reproductive system related antigen; reproductive system disorder;
 PR KW cancer; gene therapy; ds.
 PR XX
 OS Homo sapiens.
 PR XX
 PN WO200155320-A2.
 PR PN
 PD 02-AUG-2001.
 PR PD
 XX
 PF 17-JAN-2001; 2001IWO-US01339.
 PR PF
 XX 31-JAN-2000; 2000US-0179065.
 PR XX
 PR 04-FEB-2000; 2000US-0180628.
 PR PR 24-FEB-2000; 2000US-0184664.
 PR PR 02-MAR-2000; 2000US-0186350.
 PR PR 16-MAR-2000; 2000US-0189874.
 PR PR 17-MAR-2000; 2000US-0190076.
 PR PR 18-APR-2000; 2000US-0198123.

PR 19-MAY-2000; 2000US-0205515.
PR 07-JUN-2000; 2000US-0209467.
PR 28-JUN-2000; 2000US-0214886.
PR 30-JUN-2000; 2000US-0215135.
PR 07-JUL-2000; 2000US-0216647.
PR 07-JUL-2000; 2000US-0216880.
PR 11-JUL-2000; 2000US-0217487.
PR 11-JUL-2000; 2000US-0217496.
PR 14-JUL-2000; 2000US-0218290.
PR 26-JUL-2000; 2000US-0220963.
PR 26-JUL-2000; 2000US-0220964.
PR 14-AUG-2000; 2000US-0224518.
PR 14-AUG-2000; 2000US-0224519.
PR 14-AUG-2000; 2000US-0225213.
PR 14-AUG-2000; 2000US-0225214.
PR 14-AUG-2000; 2000US-0225266.
PR 14-AUG-2000; 2000US-0225267.
PR 14-AUG-2000; 2000US-0225268.
PR 14-AUG-2000; 2000US-0225270.
PR 14-AUG-2000; 2000US-0225447.
PR 14-AUG-2000; 2000US-0225757.
PR 14-AUG-2000; 2000US-0225758.
PR 14-AUG-2000; 2000US-0225759.
PR 14-AUG-2000; 2000US-0226279.
PR 18-AUG-2000; 2000US-0226681.
PR 22-AUG-2000; 2000US-0226686.
PR 22-AUG-2000; 2000US-0227182.
PR 23-AUG-2000; 2000US-0227009.
PR 30-AUG-2000; 2000US-0228924.
PR 01-SEP-2000; 2000US-0229287.
PR 01-SEP-2000; 2000US-0229343.
PR 01-SEP-2000; 2000US-0229345.
PR 01-SEP-2000; 2000US-0229345.
PR 05-SEP-2000; 2000US-0229509.
PR 05-SEP-2000; 2000US-0229513.
PR 06-SEP-2000; 2000US-0230437.
PR 06-SEP-2000; 2000US-0230439.
PR 08-SEP-2000; 2000US-0231244.
PR 08-SEP-2000; 2000US-0231244.
PR 08-SEP-2000; 2000US-0231244.
PR 08-SEP-2000; 2000US-0231414.
PR 08-SEP-2000; 2000US-0231414.
PR 08-SEP-2000; 2000US-0232080.
PR 08-SEP-2000; 2000US-0232081.
PR 12-SEP-2000; 2000US-0231968.
PR 14-SEP-2000; 2000US-0232397.
PR 14-SEP-2000; 2000US-0232399.
PR 14-SEP-2000; 2000US-0232400.
PR 14-SEP-2000; 2000US-0232401.
PR 14-SEP-2000; 2000US-0233063.
PR 14-SEP-2000; 2000US-0233064.
PR 14-SEP-2000; 2000US-0233065.
PR 21-SEP-2000; 2000US-0234223.
PR 21-SEP-2000; 2000US-0234274.
PR 25-SEP-2000; 2000US-0234997.
PR 25-SEP-2000; 2000US-0234998.
PR 26-SEP-2000; 2000US-0235834.
PR 27-SEP-2000; 2000US-0235834.
PR 27-SEP-2000; 2000US-0235836.
PR 29-SEP-2000; 2000US-0236327.
PR 29-SEP-2000; 2000US-0236367.
PR 29-SEP-2000; 2000US-0236368.
PR 29-SEP-2000; 2000US-0236369.
PR 29-SEP-2000; 2000US-0236370.
PR 02-OCT-2000; 2000US-0236802.
PR 02-OCT-2000; 2000US-0237037.
PR 02-OCT-2000; 2000US-0237039.
PR 02-OCT-2000; 2000US-0237040.
PR 13-OCT-2000; 2000US-0239935.
PR 13-OCT-2000; 2000US-0239937.
PR 20-OCT-2000; 2000US-0240960.

PR 20-OCT-2000; 2000US-0241221.
PR 20-OCT-2000; 2000US-0241785.
PR 20-OCT-2000; 2000US-0241786.
PR 20-OCT-2000; 2000US-0241787.
PR 20-OCT-2000; 2000US-0241808.
PR 20-OCT-2000; 2000US-0241809.
PR 20-OCT-2000; 2000US-0241826.
PR 01-NOV-2000; 2000US-0244617.
PR 08-NOV-2000; 2000US-0246474.
PR 08-NOV-2000; 2000US-0246475.
PR 08-NOV-2000; 2000US-0246476.
PR 08-NOV-2000; 2000US-0246477.
PR 08-NOV-2000; 2000US-0246478.
PR 08-NOV-2000; 2000US-0246523.
PR 08-NOV-2000; 2000US-0246524.
PR 08-NOV-2000; 2000US-0246525.
PR 08-NOV-2000; 2000US-0246526.
PR 08-NOV-2000; 2000US-0246527.
PR 08-NOV-2000; 2000US-0246528.
PR 08-NOV-2000; 2000US-0246532.
PR 08-NOV-2000; 2000US-0246609.
PR 08-NOV-2000; 2000US-0246610.
PR 08-NOV-2000; 2000US-0246611.
PR 08-NOV-2000; 2000US-0246613.
PR 17-NOV-2000; 2000US-0249207.
PR 17-NOV-2000; 2000US-0249208.
PR 17-NOV-2000; 2000US-0249209.
PR 17-NOV-2000; 2000US-0249210.
PR 17-NOV-2000; 2000US-0249211.
PR 17-NOV-2000; 2000US-0249212.
PR 17-NOV-2000; 2000US-0249213.
PR 17-NOV-2000; 2000US-0249214.
PR 17-NOV-2000; 2000US-0249215.
PR 17-NOV-2000; 2000US-0249216.
PR 17-NOV-2000; 2000US-0249217.
PR 17-NOV-2000; 2000US-0249218.
PR 17-NOV-2000; 2000US-0249244.
PR 17-NOV-2000; 2000US-0249245.
PR 17-NOV-2000; 2000US-0249246.
PR 17-NOV-2000; 2000US-0249264.
PR 17-NOV-2000; 2000US-0249297.
PR 17-NOV-2000; 2000US-0249300.
PR 17-NOV-2000; 2000US-0250160.
PR 01-DEC-2000; 2000US-0250391.
PR 01-DEC-2000; 2000US-0251030.
PR 05-DEC-2000; 2000US-0251988.
PR 05-DEC-2000; 2000US-0256719.
PR 06-DEC-2000; 2000US-0251479.
PR 06-DEC-2000; 2000US-0251856.
PR 08-DEC-2000; 2000US-0251868.
PR 08-DEC-2000; 2000US-0251869.
PR 08-DEC-2000; 2000US-0251899.
PR 08-DEC-2000; 2000US-0251990.
PR 11-DEC-2000; 2000US-0254097.
PR 05-JAN-2001; 2001US-0259678.
(HUMA-) HUMAN GENOME SCI INC.
Rosen CA, Barash SC, Ruben SM;
WPI; 2001-465570/50.
Isolated nucleic acid molecule encoding a reproductive system antigen
is used in preventing, treating or ameliorating a medical condition -
PS Disclosure; SEQ ID NO 9873; 1297bp + Sequence Listing: English.
xx The present invention provides the protein and coding sequences of a
CC number of human reproductive system related antigens. These can be used
CC in the prevention and treatment of reproductive system disorders,
CC including cancer. The present sequence is a genomic sequence encoding a
CC protein of the invention.

CC the cDNA. The full length clones were obtained by construction of full
CC length enriched cDNA libraries that were synthesised by the oligo-capping
CC method. The primers enable the production of the full length cDNA easily
CC without any special methods. The present sequence is a full length
CC human cDNA of the invention.
CC Note: The sequence data for this patent did not form part of the printed
CC specification, but was obtained in CD-ROM format directly from EPO.

XX Sequence 1838 BP; 404 A; 565 C; 468 G; 401 T; 0 other;

Query Match

Best Local Similarity 15.88; Score 33.2; DB 22; Length 1838;
Matches 47; Conservative 0; Mismatches 23; Indels 0; Gaps 0;

QY 59 GGGAGCTCTGGGAAAATCTCCCTCTTTGGGGGCGAGGGCGGCGGATGAGCCAGG 118
DB 816 GGGAGCAGCGGCGCAATGCGCCCTCTTGGGTGATGATGTTGTAAGGCGAGG 757
QY 119 GCGGAGAGG 128
DB 756 GCGGCGAGG 747

RESULT 10

AL36532 standard; DNA: 7601 BP.

AL36532;

08-JAN-2002 (first entry)

Human musculoskeletal system related polynucleotide SEQ ID NO 2897.

KW Cytostatic; immunosuppressive; nootropic; neuroprotective; antiviral;
KW anticancer; hepatocarcinoma; antidiabetic; antiinflammatory; antilice;
KW cardiant; gene therapy; cancer; immune disorder; cardiovascular disorder;
KW neurological disease; infection; human; secreted protein; ds.
KW musculoskeletal system; ds.

OS Homo sapiens.

PN WO200155367-A1.

PD 02-AUG-2001.

PF 17-JAN-2001; 2001WO-US01338.

XX 31-JAN-2000; 2000US-0179065.
PR 04-FEB-2000; 2000US-0180628.
PR 02-MAR-2000; 2000US-0184664.
PR 16-MAR-2000; 2000US-0186350.
PR 17-MAR-2000; 2000US-0189874.
PR 18-MAR-2000; 2000US-0190076.
PR 19-APR-2000; 2000US-0198123.
PR 07-JUN-2000; 2000US-0205515.
PR 28-JUN-2000; 2000US-0209467.
PR 30-JUN-2000; 2000US-0214886.
PR 07-JUL-2000; 2000US-0215135.
PR 07-JUL-2000; 2000US-0216647.
PR 11-JUL-2000; 2000US-0217487.
PR 11-JUL-2000; 2000US-0217496.
PR 14-JUL-2000; 2000US-0218290.
PR 26-JUL-2000; 2000US-0220963.
PR 14-AUG-2000; 2000US-0220964.
PR 14-AUG-2000; 2000US-0224518.
PR 14-AUG-2000; 2000US-0225213.
PR 14-AUG-2000; 2000US-0225214.
PR 14-AUG-2000; 2000US-0225266.
PR 14-AUG-2000; 2000US-0225267.
PR 14-AUG-2000; 2000US-0225268.

PR 14-AUG-2000; 2000US-0225270.
PR 14-AUG-2000; 2000US-0225447.
PR 14-AUG-2000; 2000US-0225757.
PR 14-AUG-2000; 2000US-0225759.
PR 18-AUG-2000; 2000US-0225759.
PR 22-AUG-2000; 2000US-0226279.
PR 22-AUG-2000; 2000US-0226681.
PR 22-AUG-2000; 2000US-0226686.
PR 23-AUG-2000; 2000US-0227182.
PR 30-AUG-2000; 2000US-0227009.
PR 01-SEP-2000; 2000US-0228924.
PR 01-SEP-2000; 2000US-0229343.
PR 01-SEP-2000; 2000US-0229344.
PR 01-SEP-2000; 2000US-0229345.
PR 05-SEP-2000; 2000US-0229509.
PR 05-SEP-2000; 2000US-0229513.
PR 06-SEP-2000; 2000US-0230437.
PR 06-SEP-2000; 2000US-0230438.
PR 08-SEP-2000; 2000US-0231242.
PR 08-SEP-2000; 2000US-0231243.
PR 08-SEP-2000; 2000US-0231244.
PR 08-SEP-2000; 2000US-0231413.
PR 08-SEP-2000; 2000US-0231414.
PR 08-SEP-2000; 2000US-0232080.
PR 12-SEP-2000; 2000US-0232081.
PR 14-SEP-2000; 2000US-0232397.
PR 14-SEP-2000; 2000US-0232398.
PR 14-SEP-2000; 2000US-0232399.
PR 14-SEP-2000; 2000US-0232400.
PR 14-SEP-2000; 2000US-0232401.
PR 14-SEP-2000; 2000US-0233063.
PR 14-SEP-2000; 2000US-0233064.
PR 21-SEP-2000; 2000US-0233065.
PR 21-SEP-2000; 2000US-0234223.
PR 25-SEP-2000; 2000US-0234274.
PR 25-SEP-2000; 2000US-0234997.
PR 26-SEP-2000; 2000US-0234998.
PR 27-SEP-2000; 2000US-0235484.
PR 27-SEP-2000; 2000US-0235834.
PR 29-SEP-2000; 2000US-0235836.
PR 29-SEP-2000; 2000US-0236327.
PR 29-SEP-2000; 2000US-0236367.
PR 29-SEP-2000; 2000US-0236368.
PR 29-SEP-2000; 2000US-0236369.
PR 29-SEP-2000; 2000US-0236370.
PR 02-OCT-2000; 2000US-0236802.
PR 02-OCT-2000; 2000US-0237037.
PR 02-OCT-2000; 2000US-0237039.
PR 02-OCT-2000; 2000US-0237039.
PR 13-OCT-2000; 2000US-0237040.
PR 13-OCT-2000; 2000US-0239935.
PR 20-OCT-2000; 2000US-0239937.
PR 20-OCT-2000; 2000US-0240960.
PR 20-OCT-2000; 2000US-0241221.
PR 20-OCT-2000; 2000US-0241785.
PR 20-OCT-2000; 2000US-0241786.
PR 20-OCT-2000; 2000US-0241787.
PR 20-OCT-2000; 2000US-0241808.
PR 20-OCT-2000; 2000US-0241809.
PR 01-NOV-2000; 2000US-0241826.
PR 08-NOV-2000; 2000US-0244617.
PR 08-NOV-2000; 2000US-0246474.
PR 08-NOV-2000; 2000US-0246475.
PR 08-NOV-2000; 2000US-0246476.
PR 08-NOV-2000; 2000US-0246477.
PR 08-NOV-2000; 2000US-0246478.
PR 08-NOV-2000; 2000US-0246523.
PR 08-NOV-2000; 2000US-0246524.
PR 08-NOV-2000; 2000US-0246525.
PR 08-NOV-2000; 2000US-0246526.
PR 08-NOV-2000; 2000US-0246527.

PR 18-AUG-2000; 2000US-0226279.
 PR 22-AUG-2000; 2000US-0226681.
 PR 22-AUG-2000; 2000US-0226688.
 PR 23-AUG-2000; 2000US-0227182.
 PR 30-AUG-2000; 2000US-0227009.
 PR 01-SEP-2000; 2000US-0228924.
 PR 01-SEP-2000; 2000US-0228287.
 PR 01-SEP-2000; 2000US-0229343.
 PR 01-SEP-2000; 2000US-0229344.
 PR 05-SEP-2000; 2000US-0229345.
 PR 05-SEP-2000; 2000US-0229513.
 PR 06-SEP-2000; 2000US-0230437.
 PR 08-SEP-2000; 2000US-0230438.
 PR 08-SEP-2000; 2000US-0231243.
 PR 08-SEP-2000; 2000US-0231244.
 PR 08-SEP-2000; 2000US-0231244.
 PR 08-SEP-2000; 2000US-0231413.
 PR 08-SEP-2000; 2000US-0231414.
 PR 08-SEP-2000; 2000US-0232080.
 PR 12-SEP-2000; 2000US-0232081.
 PR 14-SEP-2000; 2000US-0232397.
 PR 14-SEP-2000; 2000US-0232398.
 PR 14-SEP-2000; 2000US-0232399.
 PR 14-SEP-2000; 2000US-0232400.
 PR 14-SEP-2000; 2000US-0232401.
 PR 14-SEP-2000; 2000US-0233063.
 PR 14-SEP-2000; 2000US-0233064.
 PR 14-SEP-2000; 2000US-0233065.
 PR 21-SEP-2000; 2000US-0234223.
 PR 21-SEP-2000; 2000US-0234274.
 PR 25-SEP-2000; 2000US-0234997.
 PR 25-SEP-2000; 2000US-0234998.
 PR 26-SEP-2000; 2000US-0235484.
 PR 27-SEP-2000; 2000US-0235484.
 PR 27-SEP-2000; 2000US-0235834.
 PR 29-SEP-2000; 2000US-0235836.
 PR 29-SEP-2000; 2000US-0236327.
 PR 29-SEP-2000; 2000US-0236367.
 PR 29-SEP-2000; 2000US-0236368.
 PR 29-SEP-2000; 2000US-0236369.
 PR 02-OCT-2000; 2000US-0236802.
 PR 02-OCT-2000; 2000US-0237037.
 PR 02-OCT-2000; 2000US-0237038.
 PR 02-OCT-2000; 2000US-0237039.
 PR 13-OCT-2000; 2000US-0237040.
 PR 13-OCT-2000; 2000US-0239935.
 PR 20-OCT-2000; 2000US-0240960.
 PR 20-OCT-2000; 2000US-0241221.
 PR 20-OCT-2000; 2000US-0241785.
 PR 20-OCT-2000; 2000US-0241786.
 PR 20-OCT-2000; 2000US-0241787.
 PR 20-OCT-2000; 2000US-0241808.
 PR 20-OCT-2000; 2000US-0241809.
 PR 01-NOV-2000; 2000US-0241826.
 PR 08-NOV-2000; 2000US-0244617.
 PR 08-NOV-2000; 2000US-0246474.
 PR 08-NOV-2000; 2000US-0246475.
 PR 08-NOV-2000; 2000US-0246476.
 PR 08-NOV-2000; 2000US-0246477.
 PR 08-NOV-2000; 2000US-0246478.
 PR 08-NOV-2000; 2000US-0246523.
 PR 08-NOV-2000; 2000US-0246524.
 PR 08-NOV-2000; 2000US-0246525.
 PR 08-NOV-2000; 2000US-0246526.
 PR 08-NOV-2000; 2000US-0246527.
 PR 08-NOV-2000; 2000US-0246528.
 PR 08-NOV-2000; 2000US-0246532.
 PR 08-NOV-2000; 2000US-0246609.
 PR 08-NOV-2000; 2000US-0246610.
 PR 08-NOV-2000; 2000US-0246611.

PR 08-NOV-2000; 2000US-0246613.
 PR 17-NOV-2000; 2000US-0249207.
 PR 17-NOV-2000; 2000US-0249208.
 PR 17-NOV-2000; 2000US-0249209.
 PR 17-NOV-2000; 2000US-0249210.
 PR 17-NOV-2000; 2000US-0249211.
 PR 17-NOV-2000; 2000US-0249212.
 PR 17-NOV-2000; 2000US-0249213.
 PR 17-NOV-2000; 2000US-0249214.
 PR 17-NOV-2000; 2000US-0249215.
 PR 17-NOV-2000; 2000US-0249216.
 PR 17-NOV-2000; 2000US-0249217.
 PR 17-NOV-2000; 2000US-0249218.
 PR 17-NOV-2000; 2000US-0249244.
 PR 17-NOV-2000; 2000US-0249245.
 PR 17-NOV-2000; 2000US-0249264.
 PR 17-NOV-2000; 2000US-0249265.
 PR 17-NOV-2000; 2000US-0249265.
 PR 17-NOV-2000; 2000US-0249297.
 PR 17-NOV-2000; 2000US-0249299.
 PR 01-DEC-2000; 2000US-0249300.
 PR 01-DEC-2000; 2000US-0250391.
 PR 05-DEC-2000; 2000US-0250391.
 PR 05-DEC-2000; 2000US-0250398.
 PR 06-DEC-2000; 2000US-0250398.
 PR 08-DEC-2000; 2000US-0251479.
 PR 08-DEC-2000; 2000US-0251856.
 PR 08-DEC-2000; 2000US-0251856.
 PR 08-DEC-2000; 2000US-0251868.
 PR 08-DEC-2000; 2000US-0251869.
 PR 08-DEC-2000; 2000US-0251899.
 PR 11-DEC-2000; 2000US-0251990.
 PR 05-JAN-2001; 2001US-0259678.

(HUMA-) HUMAN GENOME SCI INC.
 Rosen CA, Barash SC, Ruben SM;
 WPI; 2001-451937/48.

Isolated polypeptide for treating, preventing and/or prognosing disorders related to the musculoskeletal system including musculoskeletal cancers and also for testing and detection e.g. diagnosis -

Example 2; SEQ ID NO 2698; 781bp + Sequence Listing; English.

CC The invention relates to novel genes (AAL34669-AAL37666) and proteins (AAB03087-AAB04109) associated with the musculoskeletal system useful for preventing, treating or ameliorating medical conditions e.g. by protein or gene therapy. The genes are isolated from a range of human tissues disclosed in the specification. The nucleic acids, proteins, antibodies and (ant)agonists are useful in the diagnosis, treatment and prevention of: (a) cancer, e.g. breast and ovarian cancer and other cancers of the adrenal gland, liver, lung, or urogenital; (b) gastrointestinal tract, liver, lung, or urogenital; (c) autoimmune haemolytic anaemia, autoimmune thyroiditis, diabetes mellitus, Crohn's disease, multiple sclerosis, rheumatoid arthritis and ulcerative colitis; (c) cardiovascular disorders such as myocardial ischaemias; (d) wound healing; (e) neurological diseases e.g. cerebral anoxia and epilepsy; and (f) infectious diseases such as viral, bacterial, fungal and parasitic infections.

Note: The sequence data for this patent did not form part of the printed specification, but was obtained in electronic format directly from WIPO at ftp.wipo.int/pub/published_pct_sequences.

SQ Sequence 7601 BP; 2275 A; 1450 C; 1551 G; 2325 T; 0 other;

Query Match 15.8%; Score 33.2; DB 22; Length 7601;
 Best Local Similarity 54.0%; Pred. No. 2.3;
 Matches 68; Conservative 0; Mismatches 58; Indels 0; Gaps 0;

PR 29-SEP-2000; 2000US-0236367.
 PR 29-SEP-2000; 2000US-0236368.
 PR 29-SEP-2000; 2000US-0236369.
 PR 29-SEP-2000; 2000US-0236370.
 PR 02-OCT-2000; 2000US-0236802.
 PR 02-OCT-2000; 2000US-0237037.
 PR 02-OCT-2000; 2000US-0237038.
 PR 02-OCT-2000; 2000US-0237039.
 PR 02-OCT-2000; 2000US-0237040.
 PR 13-OCT-2000; 2000US-0239335.
 PR 13-OCT-2000; 2000US-0239337.
 PR 20-OCT-2000; 2000US-0240960.
 PR 20-OCT-2000; 2000US-0241221.
 PR 20-OCT-2000; 2000US-0241785.
 PR 20-OCT-2000; 2000US-0241786.
 PR 20-OCT-2000; 2000US-0241787.
 PR 20-OCT-2000; 2000US-0241808.
 PR 20-OCT-2000; 2000US-0241809.
 PR 20-OCT-2000; 2000US-0241826.
 PR 01-NOV-2000; 2000US-0244617.
 PR 08-NOV-2000; 2000US-0246474.
 PR 08-NOV-2000; 2000US-0246475.
 PR 08-NOV-2000; 2000US-0246476.
 PR 08-NOV-2000; 2000US-0246477.
 PR 08-NOV-2000; 2000US-0246478.
 PR 08-NOV-2000; 2000US-0246523.
 PR 08-NOV-2000; 2000US-0246524.
 PR 08-NOV-2000; 2000US-0246525.
 PR 08-NOV-2000; 2000US-0246526.
 PR 08-NOV-2000; 2000US-0246527.
 PR 08-NOV-2000; 2000US-0246528.
 PR 08-NOV-2000; 2000US-0246532.
 PR 08-NOV-2000; 2000US-0246609.
 PR 08-NOV-2000; 2000US-0246610.
 PR 08-NOV-2000; 2000US-0246611.
 PR 08-NOV-2000; 2000US-0246613.
 PR 17-NOV-2000; 2000US-0249207.
 PR 17-NOV-2000; 2000US-0249208.
 PR 17-NOV-2000; 2000US-0249209.
 PR 17-NOV-2000; 2000US-0249210.
 PR 17-NOV-2000; 2000US-0249211.
 PR 17-NOV-2000; 2000US-0249212.
 PR 17-NOV-2000; 2000US-0249213.
 PR 17-NOV-2000; 2000US-0249214.
 PR 17-NOV-2000; 2000US-0249215.
 PR 17-NOV-2000; 2000US-0249216.
 PR 17-NOV-2000; 2000US-0249217.
 PR 17-NOV-2000; 2000US-0249218.
 PR 17-NOV-2000; 2000US-0249244.
 PR 17-NOV-2000; 2000US-0249245.
 PR 17-NOV-2000; 2000US-0249246.
 PR 17-NOV-2000; 2000US-0249265.
 PR 17-NOV-2000; 2000US-0249297.
 PR 17-NOV-2000; 2000US-0249299.
 PR 01-DEC-2000; 2000US-0249300.
 PR 01-DEC-2000; 2000US-0250160.
 PR 01-DEC-2000; 2000US-0250391.
 PR 05-DEC-2000; 2000US-0251030.
 PR 05-DEC-2000; 2000US-0251988.
 PR 05-DEC-2000; 2000US-0256719.
 PR 06-DEC-2000; 2000US-0251479.
 PR 08-DEC-2000; 2000US-0251856.
 PR 08-DEC-2000; 2000US-0251859.
 PR 08-DEC-2000; 2000US-0251869.
 PR 08-DEC-2000; 2000US-0251989.
 PR 11-DEC-2000; 2000US-0254097.
 PR 05-JAN-2001; 2001US-0259678.
 XX
 XX (HUMA-) HUMAN GENOME SCI INC.
 XX
 XX Rosen CA, Barash SC, Ruben SM;
 XX

DR WPI: 2001-465570/50.
 XX
 XX Isolated nucleic acid molecule encoding a reproductive system antigen
 PT is used in preventing, treating or ameliorating a medical condition -
 PS
 XX Disclosure: SEQ ID NO 7415; 1297pp + Sequence Listing; English.
 CC The present invention provides the protein and coding sequences of a
 CC number of human reproductive system related antigens. These can be used
 CC in the prevention and treatment of reproductive system disorders,
 CC including cancer. The present sequence is a genomic sequence encoding a
 CC protein of the invention.
 XX
 SQ Sequence 11360 BP; 2899 A; 2709 C; 2708 G; 3044 T; 0 other;
 Query Match 15.8%; Score 33.2; DB 22; Length 11360;
 Best Local Similarity 67.1%; Pred No. 2.6; Indels 0; Gaps 0;
 Matches 47; Conservative 0; Mismatches 23;
 QY 59 GGGAGGCTGGGAAATCTCCCTTTGGGGGCGGCGGGGCGGATGACCCAGG 118
 DB 4820 GGGAGGACGTGGGCATGGCCCTCTTGGGTGTGAATAGTGGTGTAAAGGCAAG 4761
 QY 119 GCCGAGAAG 128
 DB 4760 GCCGGAAGG 4751
 RESULT 14
 AAA94801/c
 ID AAA94801 standard; DNA; 553 BP.
 XX
 AC AAA94801;
 XX
 DT 23-FEB-2001 (first entry)
 XX
 DE Drosophila melanogaster odorant receptor DOR62 coding sequence.
 XX
 KW Odorant receptor; fruit fly; DOR62; odour recognition; pest control; ds.
 OS
 XX Drosophila melanogaster.
 FH
 FT Key Location/Qualifiers
 FT CDS 1..462
 FT /*tag= a
 FT /product= "DOR62"
 FT /partial
 PN W0200050566-A2.
 XX
 PD 31-AUG-2000.
 XX
 PE 25-FEB-2000; 2000MO-US04995.
 XX
 PR 25-FEB-1999; 99US-0257706.
 XX
 PA (UYCO) UNIV COLUMBIA NEW YORK.
 XX
 PA Vossell LB, Amrein HO, Axel R;
 XX
 PI WPI: 2000-572081/53.
 DR P-PSDB; AAB26401.
 DR
 XX Novel nucleic acid encoding an insect odorant receptor, for identifying
 PT modulator compounds that are useful in controlling pest population -
 XX
 PS Claim 4; Fig 8; 176pp; English.
 CC The present sequence is the coding sequence for the Drosophila
 CC melanogaster odorant receptor DOR62. It was isolated using a differential
 CC cloning strategy. The odorant genes and proteins are useful as they aid
 CC in the study of the olfactory organ in mammals, as well as aiding the
 CC understanding of the link between odour recognition and behaviour in

CC Insects. They also enable the identification of compounds capable of
CC activating and inhibiting the receptors, allow the control of pest
CC populations via the use of alarm odour ligands and via the use of ligands
CC which interfere with the interaction between odorant ligands and
CC receptors associated with fertility.

sequence 553 BP; 141 A; 145 C; 132 G; 135 T; 0 other,

Query Match	15.4%	Score 32.4	DB 21	Length 553
Best Local Similarity	48.4%	Pred. No. 1.8		
Matches 90	Conservative	0	Mismatches 96	Indels 0
				Gaps 0

QY	2	CGCCTTGTCTTTTCTCCGACACCGCTGCAGGAATCCCGAGAAAGCAAGCTGACTATAGGG	61
Db	264	CGCCTTCGCTCTGTCTCGCATCTGTCCCAAAATACAGATTACAAACACCTTCCAAAGT	205
QY	62	AAGCTTCGGAAATAATCTCCCTGTTTGGGGGGGCGAGCGGGGAGTAGCCAGGGCC	121
Db	204	GAGGGCGGAGAAATAATCAATTCGAGATGATCATGCGGTGTGGCTCGGTGCATCCGCTAC	145
QY	122	GAGAGGAACCTGTGAAGACTCCGTAGATTGCTCTAGACGCCCTTCAGACACTCTGGGGCA	181
Db	144	GTACACGCAATGCATGCGCAGCGGTACACTGTGAGCGCGCGGAGCAGACGCAACTTGGGCAT	85
QY	182	GGCTGC	187
Db	84	GCAAGG	79

RESULT 15

AAK92113
ID AAK92773 standard; cDNA; 555 bp.
XX

AC AAK92773;

06-NOV-2001 (first entry)

human CDNA 3'-end sequence, SEQ ID NO: 1233

XX
XX

reagents; cDNA synthesis; oligo-capping; ss.

OS Homo sapiens.

PN EP1.130094-A2.
xx

05-SEP-2001 PD YY

U/-JUL-2000; 2000EP-0114089

08-JUL-1999; 99JP-01944486
11-JAN-2000 3000TR-0118774
PB

2000, 20000P-0183/65

XX
XX
: YCMT CEN
: YCMT CEN

PI Wakamatsu A, Sugiyama

WPI; 2001-524255/58.

830 Primers useful for synthesizing full length cDNA clones and th

CC-0

The Inventor's

isolated and nucleotide sequences of 5' and 3' UTRs of cDNA molecules encoding a human protein have been

CDNA are useful for clarifying the function of the protein.

Sequence 555 BP; 124 A; 148 C; 167 G; 110 T; 6 other;
 CC without any special methods. The present sequence is the nucleotide
 CC sequence of the 3'-end of a cDNA provided in the invention.
 CC Note: The sequence data for this patent did not form part of the printed
 CC specification, but was obtained in CD-ROM format directly from EPO.
 XX
 SQ

sequence 555 BP; 124 A; 148 C; 167 G; 110 T; 6 other;

Query Match	15.4%	Score 32.4	DB 22	Length 555
Best Local Similarity	49.4%	Pred No 1.8		
Matches 84	Conservative	0	Mismatches 86	Indels 0
				Gaps 0

QY 19 TCAGCCTGTCAAGGAAGTCCAGAAAGCAGCTGACTTTAGGGAAGGTCTGGGAAAAATC 78
 Db 85 TTAAACCTGTAATTTAAAGCACACACAAACAAAGGTGTCAATGGGTAATCTCGGGGCCAACGC 144
 QY 79 TCCCTGCTTTTGGGGGGGAGGGCGGGGGGTAGCCAGAGGCGGAGAACTCTTAAG 133
 Db 145 CCCACGATGTAGGGGAGGCCGTGATTCGATCGGGGGCGGGGGCGCTTCTCCATGGGGTGC 204
 QY 139 ACTCCGTAGATTGCTTAGACCGGCTCAGACATCTCGGGGAGACGTGGA 188
 Db 205 GCCCCTTAGTGAGGTGTCACTCTCTCCAGAGTCAACAGAGGCTTTGTGGGA 254

```
Search completed: November 6, 2002, 14:17:11
Job time : 63.7914 secs
```



```

/sex="Female"
/lab_host="E. coli strain XL10-Gold, T1-resistant, F-"
/notes="Vector: PWD42nv: Purified genomic DNA from M. musculus C57Bl/6J (female) was obtained from the Jackson Laboratory Mouse DNA Resource (http://www.jax.org/resources/documents/dnares/). The DNA was hydrodynamically sheared by repeated passage through a 0.005 inch orifice at constant velocity. The sheared DNA was blunt end-repaired with T4 DNA polymerase and T4 polynucleotide kinase. Adaptor oligonucleotides were ligated to the blunt ends in high molar excess. The adaptor DNA was purified and size-selected for a 9.5 to 10.5 kb range using preparative agarose gel electrophoresis. Vector DNA was prepared from a derivative of PWD42 (91147321141gb|AF129072.1), a copy-number inducible derivative of plasmid R7. The vector was ligated with adaptors complementary to the Insert adaptors and purified. The sheared, adaptor mouse DNA was annealed to adaptor vector DNA, and transformed into chemically-competent E. coli XL10-Gold (Stratagene) cells and selected for ampicillin resistance."

```

Query Match	40.2%	Score 667.8	DB 12	Length 689
Best Local Similarity	99.7%	Pred. No. 4.2e-114		
Matches 669:	Conservative 0	Mismatches 2	Indels 0	Gaps 0

QY	924	AACCCCACTATTCTTCAGCCCTCAGATGAAAGAAAGTAATGTAACCTTGAGAGATCTGATG	983
DB	689	AACCCCACTATTCTTCAGCCCTCAGATGAAAGAAAGTAATGTAACCTTGAGAGATCTGATG	630
QY	984	GGTTCMACTGAACTAAGGCGACAGGGGTGGAGGTTTGTACCATTAACCTGAAGTGGGGTG	1043
DB	629	GGTTCMACTGAACTAAGGCGACAGGGGTGGAGGTTTGTGAACCATTAACCTGAAGTGGGGTG	570
QY	1044	TTGGTTAGTAGTGAAGCCATGAATACCATTAATATCTGTGATGGGCGCAAGATCACT	1103
DB	569	TTGGTTAGTAGTGAAGCCATGAATACCATTAATATCTGTAGTGGCGCAAGATCACT	510
QY	1104	GTCGTTCAAGAACACACAGGCCCTCCTCAGAACACAGCGGAGCAATTGAAAGGACACCACTCCG	1163
DB	509	GTCGTTCAAGAACACACAGGCCCTCCTCAGAACACAGCGGAGCAATTGAAAGGACACCACTCCG	450
QY	1164	TGCTTTCCATACCCGTTGTTTGTACCGGTGAACGCAATCACTCTCGGCATGAAACAG	1223
DB	449	TGCTTTCCATACCCGTTGTTTGTACCGGTGAACGCAATCACTCTCGGCATGAAACAG	390
QY	1224	GCTTTTCTCTGAGACCTGGGGTCTGGAAGTGTGTCTCTGAGAACAGAAACTATCTTG	1283
DB	389	GCTTTTCTCTGAGACCTGGGGTCTGGAAGTGTGTCTCTGAGAACAGAAACTATCTTG	330
QY	1284	TTACTATGAGCATTAAGTAATACACGAGACTGTGAGATAGCCCTGAGCGTGGCCGTTTAA	1343
DB	329	TTACTATGAGCATTAAGTAATACACGAGACTGTGAGATAGCCCTGAGCGTGGCCGTTTAA	270
QY	1344	AAAAGTTTGATGCTTTAGAAAGAAATCGTGCTTAAAGAAAGCCTTACCTGGCATGGGGGC	1403
DB	269	AAAAGTTTGATGCTTTAGAAAGAAATCGTGCTTAAAGAAAGCCTTACCTGGCATGGGGGC	210
QY	1404	CCATCTCTCTCAGCCATCCGAATCTCAATCTGTCGTGTGTGGTGAAGATAGATCTTCGG	1463
DB	209	CCATCTCTCTCAGCCATCCGAATCTCAATCTGTCGTGTGTGGTGAAGATAGATCTTCGG	150
QY	1464	AATGTAACCATGTCTTGCTTTTCTCTGCGGCTTGCGTGAAGGAAGTCCAGGCAAGCCTAG	1523
DB	149	AATGTAACCATGTCTTGCTTTTCTCTGCGGCTTGCGTGAAGGAAGTCCAGGCAAGCCTAG	90
QY	1524	ACGTCCTTGGGGGTAGTCTGTGGGAAATATCTCCAAAGATTTTAAAGAGGGGACAGCGGGGGA	1583
DB	89	ACGTCCTTGGGGGTAGTCTGTGGGAAATATCTCCAAAGATTTTAAAGAGGGGACAGCGGGGGA	30
QY	1584	TGACAAACTTG	1594

Db 29 TGAGAACTCG 19

RESULT 2
CNS03SBW
LOCUS
DEFINITION

Accession	Size	Library	Survey
CNS03SBW	946 bp	Linear	GSS 17-MAY-2000
Tetraodon nigroviridis genome survey sequence T7 end of clone			
U51620 of library G from Tetraodon nigroviridis, genomic survey sequence.			

ACCESSION	AL258341
VERSION	AL258341.1
KEYWORDS	GSS; genome survey sequence
SOURCE	Tetraodon nigroviridis.
ORGANISM	Tetraodon nigroviridis

REFERENCE
AUTHORS

Roest-Crolius, H., Jaillon, O., Dasilva, C., Fitzames, C., Fisher, C., Bouneau, L., Billault, A., Quetier, F., Saurin, W., Bernot, A. and Weissenbach, J.

TITLE
Characterization and repeat analysis of the compact genome of the freshwater pufferfish *Tetraodon nigroviridis*
JOURNAL
Unpublished
REFERENCE
2 (bases 1 to 946)
AUTHORS
Roest-Crolius, H., Jallion, O., Dasilva, C., Bouneau, L., Fisher, C.

TITLE

Human gene number estimate provided by genome wide analysis using Tetraodon nigroviridis DNA sequence

JOURNAL
REFERENCE
AUTHORS
TITLE
JOURNAL
COMMENT

Submitted (12-APR-2000) to the EMBL/GenBank/DBJ databases. This sequence is a single read and was generated as part of a large scale clone-end sequencing project of the Tetraodon nigriviridis genome. For more information, please take a look at <http://www.genoscope.cns.fr/Tetraodon>.

FEATURES

1. .946

BASE COUNT	/note="Genoscope sequence ID : C08C051BD10LP1-end : T7"			
ORIGIN	191 a	291 c	305 g	133 t
				26 others

Query Match	4.2%	Score 70.4	DB 12	Length 946
Best Local Similarity	67.9%	Pred. No. 5.5e-05		
Matches 95; Conservative	2;	Mismatches 43;	Indels 0;	Gaps 0;

155 GGTCCTTCTCCCCACCAATGGACCGGGAGCTCCACTGATATGGACAGATAGCTTTA 214

215 CAGCTACATTCAAAACACACACACACACACACACACACACACACACAC 274

275 ACACATGTTTCTTCCCTCC 294

Db 827 ACTCATGAGGTATCCCTCC 846

RESULT 3

LOCUS	CNS02GTW	681 bp	DNA	linear	GSS 13-MAY-2000
DEFINITION	Tetradon nigroviridis genome survey sequence PUC-Orl end of clone				

138G22 of library G from *Tetraodon nigroviridis*, genomic survey

ACCESSION	VERSION	KEYWORDS	SOURCE	ORGANISM
AL196781	GI:7834931	GSS, genome survey sequence.	Tetraodon nigroviridis.	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Actinopterygii; Neopterygii; Teleostei; Euteleostei; Neoteleostei; Acanthomorpha; Acanthopterygii; Percormorpha; Tetraodontiformes; Tetraodontidae; Tetraodon.
REFERENCE	AUTHORS	TITLE	JOURNAL	REFERENCE
1 (bases 1 to 681)	Ruest-Crollius,H., Jallion,O., Dasilva,C., Filames,C., Fisher,C., Bouneau,L., Billault,A., Queller,F., Saurin,W., Bernot,A. and Weissenbach,J.	Characterization and repeat analysis of the compact genome of the freshwater pufferfish Tetraodon nigroviridis	Unpublished	2 (bases 1 to 681)
3 (bases 1 to 681)	Saurin,W. and Weissenbach,J.	Human gene number estimate provided by genome wide analysis using Tetraodon nigroviridis DNA sequence	Unpublished	3 (bases 1 to 681)
Genoscope.	Direct Submission	Submitted (12-APR-2000) to the EMBL/GenBank/DBJ databases	This sequence is a single read and was generated as part of a large scale clone-and-sequencing project of the Tetraodon nigroviridis genome. For more information, please take a look at http://www.genoscope.cns.fr/tetraodon .	Location/Qualifiers
1. 681	/organism="Tetraodon nigroviridis"	/db_xref="taxon:99883"	/clone="138G22"	/clone.lib="G"
/note="Genoscope sequence ID : COAG138BD11SP1-end : PUC-Orl"				
BASE COUNT	199 a	121 c	169 g	166 t
ORIGIN	26 others			
Query Match	4.0%	Score 66.8	DB 12	Length 681
Matches	Similarity 81.9%	Pred. NO. 0.00033	Matches 77	Conservative 0
Mismatches	17	Indels 0	Gaps 0	
Qy	229	ACATGTTTCTT	288	
Db	569	ACATGTTGCT	530	
Qy	289	CCCTCCCTCCACTTCTCCCATTCCTGTGCTC	322	
Db	529	CCCTCCCTCCCTCCCTCATTCCTTTTGTTC	496	
RESULT 4	A2463187	511 bp	DNA	linear
LOCUS	A2463187	10kb plasmid	U06C1M	library Mus musculus genomic
DEFINITION	clone U06C1M0271B21 R, DNA sequence.			
ACCESSION	A2463187			
VERSION	A2463187.1			
KEYWORDS	GSS.			
SOURCE	house mouse.			
ORGANISM	Mus musculus			
REFERENCE	AUTHORS	TITLE	JOURNAL	REFERENCE
1 (bases 1 to 511)	Dunn,D., Aoyagi,A., Barber,M., Beacorn,T., Duval,B., Hamil,C., Islami,H., Longacre,S., Mahmoud,M., Meenen,E., Pedersen,T., Reilly,M., Rose,M., Rose,R., Stokes,R., Tingey,A., von Niederhausern,A. and Wright,D., Weiss,R.			

TITLE	JOURNAL	COMMENT
Mouse whole genome scaffolding with paired end reads from 10kb plasmid inserts	Unpublished (2000)	Contact: Robert B. Weiss University of Utah Rm. 308, Biomedical Polymers Research Bldg., 20 S. 2030 E., SLIC, UT 84112, USA Tel: 801 585 5606 Fax: 801 585 7177 Email: rdunn@genetics.utah.edu Insert Length: 10000 Std Error: 0.00 Plate: 0271 row: B column: 21 Seq primer: CACACGAGAAACAGCATATGACC Class: plasmid ends High quality sequence stop: 511.
FEATURES	source	Location/Qualifiers
	1. 511	/organism="Mus musculus" /strain="C57BL/6J" /db.xref="taxon:10090" /clone="U06C1M0271B21" /clone_11b="Mouse 10kb plasmid U06C1M library" /sex="Male" /lab_host="E. Coli strain XL10-Gold, T1-resistant, F-" /note="Vector: pMD42nv; Purified genomic DNA from the Jackson Laboratory Mouse DNA Resource (http://www.jax.org/resources/documents/dnares/). The DNA was hydrodynamically sheared by repeated passage through a 0.005 inch orifice at constant velocity. The sheared DNA was blunt end-repaired with T4 DNA polymerase and T4 polynucleotide kinase. Adaptor oligonucleotides were ligated to the blunt ends in high molar excess. The adapted DNA was purified and size-selected for a 9.5 to 10.5 kb range using preparative agarose gel electrophoresis. Vector DNA was prepared from a derivative of pMD42 (g11473211419b1aF129072.1) a copy-number inducible derivative of plasmid R1. The vector was ligated with adaptors complementary to the insert adaptors and purified. The sheared, adapted mouse DNA was annealed to adapted vector DNA, and transformed into chemically-competent E. coli XL10-Gold (Stratagene) cells and selected for ampicillin resistance."
BASE COUNT	139 a 165 c 78 g 129 t	
ORIGIN		
Query Match	4.0%; Score 66.2; DB 12; Length 511;	
Best Local Similarity	68.1%; Pred. No. 0.00045;	
Matches 92; Conservative 0; Mismatches 43; Indels 0; Gaps 0;		
QY 181	CGGAGCTCCACGTATATGACAGATAGCTTACAGCTACATTCMAAACACACACAC 240	
Db 246	CAGGAGAGATAGCTTATGTTCTTCTGCGCTTCCACACAACATCACTACACACAC 305	
QY 241	ACACACACACACACACACACACACACACACATGTTTCTTCCTCCCTCCAC 300	
Db 306	ACACACACACACACACACACACACACACACACATGAGCCATCATCACTCTCC 365	
QY 301	TTCCTCCATCTCTC 315	
Db 366	TGCTTCGCGTTCTC 380	
RESULT 5		
CNS038VZ/c	958 bp DNA linear GSS 15-MAY-2000	
LOCUS		
DEFINITION	Tetradon nigroviridis genome survey sequence T7 end of clone 006A06 of library G from Tetradon nigroviridis, genomic survey sequence.	
ACCESSION	AL233144	
VERSION	AL233144.1 GI:7892279	
KEYWORDS	GSS; genome survey sequence.	

SOURCE Tetraodon nigroviridis.
ORGANISM Tetraodon nigroviridis
Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi; Actinopterygii; Neopterygii; Teleostei; Euteleostei; Acanthomorpha; Acanthopterygii; Percomorpha; Tetraodontiformes; Tetraodontidae; Tetraodon.
REFERENCE 1 (bases 1 to 958)
AUTHORS Roest-Crolius,H., Jalllon,O., Dasilva,C., Fizames,C., Fisher,C., Bouneau,L., Billault,A., Queller,F., Saurin,W., Bernot,A. and Weissenbach,J.
TITLE Characterization and repeat analysis of the compact genome of the freshwater pufferfish Tetraodon nigroviridis
JOURNAL Unpublished
REFERENCE 2 (bases 1 to 958)
AUTHORS Roest-Crolius,H., Jalllon,O., Dasilva,C., Bouneau,L., Fisher,C., Saurin,W., Fizames,C., Wincker,P., Brottier,P., Queller,F., Bernot,A. and Weissenbach,J.
TITLE Human gene number estimate provided by genome wide analysis using Tetraodon nigroviridis DNA sequence
JOURNAL Unpublished
REFERENCE 3 (bases 1 to 958)
AUTHORS Direct Submission
TITLE Submitted (12-APR-2000) to the EMBL/GenBank/DBJ databases
COMMENT This sequence is a single read and was generated as part of a large scale clone-and-sequencing project of the Tetraodon nigroviridis genome. For more information, please take a look at <http://www.genoscope.cns.fr/Tetraodon>.
FEATURES
source
1..958
/organism="Tetraodon nigroviridis"
/db_xref="taxon:99883"
/clone="006A06"
/clone_1lb="G"
/note="Genoscope sequence ID : C05G006BA03LPI-end : T7"
BASE COUNT 247 a 202 c 229 g 276 t 4 others
ORIGIN
Query Match 3.9%; Score 65.4; DB 12; Length 958;
Best Local Similarity 63.9%; Pred. No. 0.00066;
Matches 99; Conservative 0; Mismatches 56; Indels 0; Gaps 0;
QY 131 AAGATGGAATATTTGTACAAAGAGGCTTTCTCCGCCACATCGACCCGGAGCTCC 190
Db 701 ATGAAAGAAATCTACGATGACAGTTGATACCTGTCAGAGGGGAGCAGCAGCGGAC 642
QY 191 ACATGATATGACACAAATAGCTTTACAGCTACATTCAAAACACACACACACACAC 250
Db 641 ACAGATTAACCCACAGATAACGACAGCTTGATATACACACACACACACACACAC 582
QY 251 ACACACACACACACACACACACACACATGTTT 285
Db 581 ACACACACACACACACACACACACACACGCTT 547

RESULT 6
LOCUS A2394294 684 bp DNA linear GSS 03-OCT-2000
DEFINITION 1M0157J18R Mouse 10kb plasmid UUGC1M library Mus musculus genomic
ACCESSION A2394294
VERSION A2394294.1 GI:10509366
KEYWORDS GSS.
SOURCE house mouse.
ORGANISM Mus musculus
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
REFERENCE 1 (bases 1 to 684)
AUTHORS Dunn,D., Aoyagi,A., Barber,M., Beacorn,T., Duval,B., Hamill,C., Islam,H., Longacre,S., Mahmoud,M., Meenen,E., Pedersen,T., Reilly,M., Rose,M., Rose,R., Stokes,R., Tingey,A., von Niederhausern,A. and Wright,D., Weiss,R.
TITLE Mouse whole genome scaffolding with paired end reads from 10kb

JOURNAL Plasmid inserts
COMMENT Unpublished (2000)
CONTACT: Robert B. Weiss
UNIVERSITY OF UTAH Genome Center
UNIVERSITY OF UTAH
Rm. 308, Biomedical Polymers Research Bldg., 20 S. 2030 E., SLIC, UT 84112, USA
TEL: 801 585 5606
FAX: 801 585 7177
EMAIL: ddunn@genetics.utah.edu
INSERT LENGTH: 10000 Std Error: 0.00
PLATE: 0157 row: 7 column: 18
SEG PRIMER: CACACAGAAACACGATGACCC
CLASS: plasmid ends
HIGH QUALITY SEQUENCE STOP: 684.
FEATURES
source
1..684
/organism="Mus musculus"
/strain="C57BL/6J"
/db_xref="taxon:10090"
/clone="UUGC1M0157J18"
/clone_1lb="Mouse 10kb plasmid UUGC1M library"
/sex="Male"
/lab_host="E. coli strain XL10-Gold, T1-resistant, F-"
/note="Vector: PMD42ny; Purified genomic DNA from M. musculus C57BL/6J (male) was obtained from the Jackson Laboratory Mouse DNA Resource
(<http://www.jax.org/resources/documents/dnares/>). The DNA was hydrodynamically sheared by repeated passage through a 0.005 inch orifice at constant velocity. The sheared DNA was blunt end-repaired with T4 DNA polymerase and T4 polynucleotide kinase. Adaptor oligonucleotides were ligated to the blunt ends in high molar excess. The adapted DNA was purified and size-selected for a 9.5 to 10.5 kb range using preparative agarose gel electrophoresis. Vector DNA was prepared from a derivative of pMD42 (g14732144bp/AF129072.1), a copy-number inducible derivative of plasmid R1. The vector was ligated with adaptors complementary to the insert adaptors and purified. The sheared, adapted mouse DNA was annealed to adaptor vector DNA, and transformed into chemically-competent E. coli XL10-Gold (Stratagene) cells and selected for ampicillin resistance."
BASE COUNT 212 a 159 c 133 g 180 t
ORIGIN
Query Match 3.9%; Score 65.2; DB 12; Length 684;
Best Local Similarity 69.8%; Pred. No. 0.00073;
Matches 88; Conservative 0; Mismatches 38; Indels 0; Gaps 0;
QY 155 GGTGCTTTCCTCCGCCACATGCGAGCGGAGCTGATATGACAGAAATAGCTTAA 214
Db 145 GGTTCATATCTCTGTACAAATAACTGGGACATACAGACTTGTCTCAAAAGCTTAA 204
QY 215 CAGCTACATTCAAAC 274
Db 205 AAGGTCATTAAC 264
QY 275 ACACAT 280
Db 265 ACACCT 270

RESULT 7
LOCUS CNS020RH 729 bp DNA linear GSS 12-MAY-2000
DEFINITION Tetraodon nigroviridis genome survey sequence T7 end of clone 224G15 of library G from Tetraodon nigroviridis, genomic survey
ACCESSION CNS020RH
VERSION AL175958
KEYWORDS GSS; genome survey sequence.
SOURCE Tetraodon nigroviridis.

QY 287 TTCCCTCCCTCCACT 301
||||| | | |||
Db 423 TTCCGACAAATGTRACT 409

Search completed: November 6, 2002, 16:06:59
Job time : 2070.91 secs

GenCore version 5.1.3
Copyright (c) 1993 - 2002 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: November 6, 2002, 14:20:41 ; Search time 285.578 Seconds
(without alignments) 15388.369 Million cell updates/sec

Title: US-09-833-740-7_COPY_1_210

Perfect score: 210
Sequence: 1 ccgcctgtttcttctctc.....ggatttgcaacattcc 210

Scoring table: IDENTITY_NUC
Gapop 10.0 , Gapext 1.0

Searched: 1797656 seqs, 10463268293 residues

Total number of hits satisfying chosen parameters: 3595312

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%

Listing first 45 summaries

Database :

GenEmbl:*
1: gb_ba:*
2: gb_htg:*
3: gb_in:*
4: gb_om:*
5: gb_ov:*
6: gb_pat:*
7: gb_ph:*
8: gb_pl:*
9: gb_pr:*
10: gb_ro:*
11: gb_sts:*
12: gb_sy:*
13: gb_un:*
14: gb_vi:*
15: em_ba:*
16: em_fun:*
17: em_hum:*
18: em_in:*
19: em_mu:*
20: em_om:*
21: em_or:*
22: em_ov:*
23: em_pat:*
24: em_ph:*
25: em_pl:*
26: em_ro:*
27: em_sts:*
28: em_un:*
29: em_vi:*
30: em_htg_hum:*
31: em_htg_inv:*
32: em_htg_other:*
33: em_htgo_inv:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
------------	-------	-------------	--------	-------	-------------

1	210	100.0	527	6	AX354813	AX354813 Sequence
2	210	100.0	198271	2	AC026591	AC026591 Homo sapi
3	210	100.0	210719	2	AC087646	AC087646 Homo sapi
4	210	100.0	213359	2	AC069006	AC069006 Homo sapi
5	62.2	29.6	2170	6	AX354807	AX354807 Sequence
6	62.2	29.6	3130	10	AF338224	AF338224 Mus muscu
7	62.2	29.6	207160	2	AC016464	AC016464 Mus muscu
8	62.2	29.6	211173	2	AL646097	AL646097 Mus muscu
9	54	25.7	493	6	AX354812	AX354812 Sequence
10	38.6	18.4	1612	2	AK021842	AK021842 Homo sapi
11	38.6	18.4	202682	2	AL513327	AL513327 Homo sapi
12	36.8	17.5	147868	2	AC106475	AC106475 Rattus no
13	35.8	17.0	145617	2	AC096681	AC096681 Canis fam
14	35.8	17.0	148198	2	AC024232	AC024232 Homo sapi
15	35.8	17.0	176831	2	AC084337	AC084337 Homo sapi
16	35.6	17.0	153394	2	AC092731	AC092731 Felis cat
17	35.2	16.8	125020	2	AF429315	AF429315 Homo sapi
18	35.2	16.8	203912	9	AC008755	AC008755 Homo sapi
19	35	16.7	166458	2	AC084737	AC084737 Homo sapi
20	34.6	16.5	105288	9	AC012330	AC012330 Homo sapi
21	34.6	16.5	145598	9	AC008132	AC008132 Homo sapi
22	34.6	16.5	150724	9	AP000550	AP000550 Homo sapi
23	34.6	16.5	150754	9	AC023491	AC023491 Homo sapi
24	34.6	16.5	159550	2	AC013360	AC013360 Homo sapi
25	34.6	16.5	163795	9	AP000356	AP000356 Homo sapi
26	34.6	16.5	165050	9	AC007325	AC007325 Homo sapi
27	34.6	16.5	170667	2	AC093573	AC093573 Pan trogl
28	34.6	16.5	175358	9	AC007981	AC007981 Homo sapi
29	34.6	16.5	180884	9	AC008018	AC008018 Homo sapi
30	34.2	16.3	64836	2	AC102520	AC102520 Mus muscu
31	34.2	16.3	191656	2	AC068733	AC068733 Homo sapi
32	34.2	16.3	214194	2	AC073757	AC073757 Mus muscu
33	34.2	16.3	233448	2	AC073692	AC073692 Mus muscu
34	34.2	16.3	237613	2	AC025769	AC025769 Homo sapi
35	34	16.2	197426	3	CNS07EG9	AL590442 Chromosom
36	33.8	16.1	148054	9	AC026467	AC026467 Homo sapi
37	33.6	16.0	141990	9	AC004691	AC004691 Homo sapi
38	33.4	15.9	1368	1	PSEFUDH	D32042 Pseudomonas
39	33.4	15.9	62875	2	AC068436	AC068436 Homo sapi
40	33.4	15.9	110000	2	HSS171M.2	Continuation (3 of
41	33.4	15.9	118229	2	AC013315	AC013315 Homo sapi
42	33.4	15.9	148539	2	AC094804	AC094804 Rattus no
43	33.4	15.9	177990	2	AC016385	AC016385 Homo sapi
44	33.4	15.9	205802	2	AC015802	AC015802 Homo sapi
45	33.4	15.9	340000	9	HS21C102	AL163302 Homo sapi

ALIGNMENTS

RESULT 1	AX354813	527 bp	DNA	Linear	PAT 06-FEB-2002
LOCUS	AX354813	Sequence	7 from Patent WO0179290.		
DEFINITION	AX354813				
ACCESSION	AX354813.1	GI:18619545			
VERSION					
KEYWORDS					
SOURCE					
ORGANISM					

human.
Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
REFERENCE
1 (sites)
AUTHORS
Drucker/D.J. and Lovshin,J.A.
TITLE
Glp-2 receptor gene promoter and uses thereof
JOURNAL
Patent: WO 0179290-A 7 25-Oct-2001;
1149336 ONTARIO INC. (CA)
FEATURES
Location/Qualifiers
source
1..527
/organism="Homo sapiens"
/db_xref="taxon:9606"

BASE COUNT	111 a	140 c	163 g	113 t
------------	-------	-------	-------	-------

	Query Match	100.0%	Score 210;	DB 6;	Length 527;
	Best Local Similarity	100.0%	Pred. NO. 7.7e-54;		
	Matches 210;	Conservative 0;	Mismatches 0;	Indels 0;	Gaps 0;
QY	1	CCGCGCTGTTCTTCTCTCCTCAGCCCTGTCAAGGAAGTCCAGAAAGACACAGCTGACTTAGG	60		
Db	1	CCGCGCTGTTCTTCTCTCCTCAGCCCTGTCAAGGAAGTCCAGAAAGACACAGCTGACTTAGG	60		
QY	61	GAAAGTCTGGGAAAAAATCTCCCTGCTTTTGGGGGGGACAGGGCGGGGGATGAGCCAGGGC	120		
Db	61	GAAAGTCTGGGAAAAAATCTCCCTGCTTTTGGGGGGGACAGGGCGGGGGATGAGCCAGGGC	120		
QY	121	CGAAGAGCAACTCTGAACACTCCGCTAGATTGCTCTAGACCGCCTCAGACACTCTCGCGGC	180		
Db	121	CGAAGAGCAACTCTGAAGACTCCGCTAGATTGCTCTAGACCGCCTCAGACACTCTCGCGGC	180		
QY	181	AGCGTGCAGAGCAATTTGTGCAAAACATTGCC	210		
Db	181	AGCGTGCAGAGCAATTTGTGCAAAACATTGCC	210		

RESULT 2	AC026591/c	LOCUS	DEFINITION	AC026591	198271 bp	DNA	linear	HTG 27-AUG-2001
			Homo sapiens chromosome 17 clone RP11-655d3 map 17, *** SEQUENCING IN PROGRESS ***, 4 ordered pieces.	AC026591				
ACCESSION				AC026591				
VERSION				AC026591.9	GI:15291083			
KEYWORDS				HTG; HTGS_PHASE2; HTGS_FULLTOP; HTGS_ACTIVEFIN.				
SOURCE				human.				
ORGANISM				Homo sapiens				

REFERENCE	1 (bases 1 to 198271)
AUTHORS	Birtten,B., Linton,L., Nusbaum,C. and Lander,E.
TITLE	Homo sapiens chromosome 17, clone RP11-655D3
JOURNAL	Unpublished
REFERENCE	2 (bases 1 to 198271)
AUTHORS	Birtten,B., Linton,L., Nusbaum,C., Lander,E., Abraham,H., Allen,N.

TITLE	Direct Submission
JOURNAL	Submitted (22-MAR-2000) Whitehead Institute/MIT Center for Genome Research, 320 Charles Street, Cambridge, MA 02141, USA
COMMENT	On Aug 26, 2001 this sequence version replaced at:14547870.

Center: Whitehead Institute/ MIT Center for Genome Research
Center code: WIBR
Web site: <http://www-seq.wi.mit.edu>
Contact: sequence.submissions@genome.wi.mit.edu
----- Project Information
Center project name: L8450

```

-----
Center clone name: 655_D_3
*
* NOTE: This is a 'working draft' sequence. It currently
* consists of 4 contigs. Gaps between the contigs
* are represented as runs of N. The order of the pieces
* is believed to be correct as given, however the sizes
* of the gaps between them are based on estimates that have
* provided by the submittor.
* This sequence will be replaced
* by the finished sequence as soon as it is available and
* the accession number will be preserved.
*
* 1 153566: contig of 153566 bp in length
* 153567 153666: gap of 100 bp
* 153667 158888: contig of 5222 bp in length
* 158889 158988: gap of 100 bp
* 158989 189474: contig of 30486 bp in length
* 189475 189574: gap of 100 bp
* 189575 198271: contig of 8657 bp in length.
*
* Location/Qualifiers
*   1..198271
*     /organism="Homo sapiens"
*     /db_xref="taxon:9606"
*     /chromosome="17"
*     /map="17"
*     /clone="RP11-655D3"
*     /clone_lib="RC11-11 Human Male BAC"
BASE COUNT      52858 a 45543 c 46237 g 53270 t      363 others
ORIGIN

```

	Query Match	100.0%	Score 210:	DB 21:	Length 198271:
	Best Local Similarity	100.0%	Pred. No. 7,5e-54:		
	Matches	210:	Conservative	0:	Mismatches 0: Indels 0: Gaps 0:
QY	1	CCGCTTTCTTCTTCTCTCA	CGCTGTCAAGGAAGTCCCGA	AAACACAGCTAGCTTAGG	60
Db	144273	CCGCTTTCTTCTTCTCTCA	CGCTGTCAAGGAATCCCGA	AAACACAGCTAGCTTAGG	144214
QY	61	GAAAGTCGGGAAAAAAT	CCCTCGTTTGGGGGGGCA	AGGGCGGGGGATGAGCCAGGGC	120
Db	144213	GAAAGTCGGGAAAAAAT	CCCTCGTTTGGGGGGGCA	AGGGCGGGGGATGAGCCAGGGC	144154
QY	121	CGAAGAGGAACCTG	GAAGACTCCGTAGATTGCT	TAGACCGGCTCAGACACTCTCGGGC	180
Db	144153	CGAAGAGGAACCTG	GAAGACTCCGTAGATTGCT	TAGACCGGCTCAGACACTCTCGGGC	144094
QY	181	AGCGTGAGAGAGATT	GTGCAAAACATTCC	210	
Db	144093	AGCGTGAGAGAGATT	GTGCAAAACATTCC	144064	

[illegible]

Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
1 (bases 1 to 210719)
Birren, B., Linton, L., Nusbaum, C., and Lander, E.
Homo sapiens chromosome 17, clone RP11-773021
Unpublished
2 (bases 1 to 210719)
Birren, B., Linton, L., Nusbaum, C., Lander, E., Allen, N., Anderson, S., Barra, N., Bastien, V., Boguslavsky, L., Bouckhvalter, B., Brown, A., Camarata, J., Campplano, A., Chepel, Y., Colangelo, M., Collins, S., Collymore, A., Cooke, P., Deatellano, K., Dewar, K., Diaz, J. S., Dodge, S., Faro, S., Ferreira, P., Fitzhugh, W., Gage, D., Galagan, J., Gardyna, S., Glade, S., Goyette, M., Graham, L., Grand-Pierre, N.,

TITLE
JOURNAL
COMMENT

Hagos, B., Heatford, A., Horton, L., Hulme, W., Iliev, I., Johnson, R., Jones, C., Karatas, A., LaRoque, K., Lamazares, R., Landers, T., Lehoczy, J., Levine, R., Liu, G., Maclean, C., MacDonald, P., Marquis, N., Matthews, C., McCarthy, M., McEwan, P., McKernan, K., McPheeters, R., Meldrim, J., Menus, L., Mihova, T., Mlenga, V., Murphy, T., Naylor, J., Nguyen, C., Norbu, C., Norman, C. H., O'Connor, T., O'Donnell, P., O'Neill, D., Oliver, J., Peterson, K., Phunkhang, P., Pierre, N., Pollara, V., Raymond, C., Retta, R., Rieback, M., Riley, R., Rise, C., Rogov, P., Roman, J., Rosetti, M., Roy, A., Santos, R., Schauer, S., Schupbach, R., Seaman, S., Severy, P., Sougne, C., Spencer, B., Stange-Thomann, N., Stojanovic, N., Strauss, M., Subramanian, A., Talamas, J., Testaye, S., Theodore, J., Travers, N., Travis, N., Trigglio, J., Vassiliev, H., Viel, R., Vo, A., Wilson, B., Wu, X., Wyman, D., Ye, W. J., Young, G., Zainoun, J., Zembek, L., Zimmer, A. and Zody, M.

Direct Submission
Submitted (15-JAN-2001) Whitehead Institute/MIT Center for Genome Research, 320 Charles Street, Cambridge, MA 02141, USA
On May 31, 2001 this sequence version replaced g1:13493095.
All repeats were identified using RepeatMasker:
Smit, A.F.A. & Green, P. (1996-1997)
<http://ftp.genome.washington.edu/RM/RepeatMasker.html>

Genome Center

Center: Whitehead Institute/ MIT Center for Genome Research

Center code: WIBR

Web site: <http://www-seq.wi.mit.edu>

Contact: sequence_submissions@genome.wi.mit.edu

Project Information

Center Project name: L11995

Center Clone name: 773_O_21

Summary Statistics

Sequencing vector: Plasmid; n/a; 100% of reads
Chemistry: Dye-terminator Big Dye, 100% of reads
Assembly program: Phrap, version 0.960731
Consensus quality: 205962 bases at least Q40
Consensus quality: 208431 bases at least Q30
Consensus quality: 209109 bases at least Q20
Insert size: 205000; agarose-fp
Insert size: 209619; sum-of-contigs
Quality coverage: 8.4 in Q20 bases; agarose-fp
Quality coverage: 8.2 in Q20 bases; sum-of-contigs

* NOTE: This is a 'working draft' sequence. It currently
* consists of 12 contigs. The true order of the pieces
* is not known and their order in this sequence record is
* arbitrary. Gaps between the contigs are represented as
* runs of N, but the exact sizes of the gaps are unknown.
* This record will be updated with the finished sequence
* as soon as it is available and the accession number will
* be preserved.

75254: contig of 75254 bp in length
75255 75354: gap of 100 bp
75355 76881: contig of 1527 bp in length
76882 76981: gap of 100 bp
76982 78538: contig of 1557 bp in length
78539 78638: gap of 100 bp
78639 79946: contig of 1308 bp in length
79947 80046: gap of 100 bp
80047 82098: contig of 2052 bp in length
82099 82198: gap of 100 bp
82199 86309: contig of 4111 bp in length
86310 86409: gap of 100 bp
86410 89601: contig of 3192 bp in length
89602 89701: gap of 100 bp
89702 105000: contig of 15299 bp in length
105001 105100: gap of 100 bp
105101 121101: contig of 16001 bp in length
121102 121201: gap of 100 bp
121202 154071: contig of 32870 bp in length
154072 154171: gap of 100 bp
154172 193985: contig of 39814 bp in length
193986 194085: gap of 100 bp
194086 210719: contig of 16634 bp in length.

FEATURES
source Location/Qualifiers

1..210719
/organism="Homo sapiens"
/db_xref="taxon:9606"
/chromosome="17"
/map="17"
/clone="RP11-773021"
/clone_11b="RPC1-11 Human Male BAC"
1..75254
/note="assembly-fragment
clone_end:SP6
vector_side:left"
75355..76881
/note="assembly-fragment"
76982..78538
/note="assembly-fragment"
78639..79946
/note="assembly-fragment"
80047..82098
/note="assembly-fragment"
82199..86309
/note="assembly-fragment"
86410..89601
/note="assembly-fragment"
89702..105000
/note="assembly-fragment"
105101..121101
/note="assembly-fragment"
121202..154071
/note="assembly-fragment"
154172..193985
/note="assembly-fragment"
194086..210719
/note="assembly-fragment
clone_end:T7
vector_side:right"

BASE COUNT 58923 a 47911 c 46884 g 55896 t 1105 others
ORIGIN

Query Match 100.0%; Score 210; DB 2; Length 210719;
Best Local Similarity 100.0%; Pred. NO. 7.5e-54;
Matches 210; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 CCGCTTCTTCTTCTCTCCTCAGCTGTCAAGAGAAATCCAGAAACACAGCTGACTTGG 60
Db 72432 CCGCTTCTTCTTCTCTCCTCAGCTGTCAAGAAATCCAGAAACACAGCTGACTTGG 72373
OY 61 GAAGCTGTGGAAAAATCTCCCTGCTTTGGGGGGGCGAGGGGCGGAGTGAAGCCAGGCG 120
Db 72372 GAAGCTGTGGAAAAATCTCCCTGCTTTGGGGGGGCGAGGGGCGGAGTGAAGCCAGGCG 72313
OY 121 CGAGAAGAACTCTGAACACCCGAGATGCTGTAGACCCGCTCAGACACTCGGGCG 180
Db 72312 CGAGAAGAACTCTGAACACCCGAGATGCTGTAGACCCGCTCAGACACTCGGGCG 72253
OY 181 AGCGTGAGAGAGATTGTGCAAAACATTTC 210
Db 72252 AGCGTGAGAGAGATTGTGCAAAACATTTC 72223

RESULT 4 21359 bp DNA linear HTG 11-DEC-2000
AC069006/c
LOCUS Homo sapiens chromosome RPCI-11 clone RP11-773021, WORKING DRAFT
DEFINITION SEQUENCE, 16 unordered pieces.
ACCESSION AC069006
VERSION AC069006.4 GI:11612641
KEYWORDS HTG; HTGS_PHASE1; HTGS_DRAFT.
SOURCE human.
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
REFERENCE 1 (bases 1 to 213359)

AUTHORS Waterston, R.H.
TITLE The sequence of Homo sapiens clone
JOURNAL Unpublished
REFERENCE 2 (bases 1 to 213359)
AUTHORS Waterston, R.H.
TITLE Direct Submission
JOURNAL Submitted (16-MAY-2000) Genome Sequencing Center, Washington
University School of Medicine, 4444 Forest Park Parkway, St. Louis,
MO 63108, USA
On Dec 11, 2000 this sequence version replaced gi:9280811.

----- Genome Center -----
Center: Washington University Genome Sequencing Center
Center code: WUGSC
Web site: http://genome.wustl.edu/gsc/index.shtml
----- Project Information -----
Center project name: H_NH0773021
----- Summary Statistics -----
Sequencing vector: M13; 100%
Chemistry: Dye-terminator Big Dye; 0% of reads
Chemistry: Dye-terminator Big Dye; 0% of reads
Assembly program: Phrap; version 0.990319
Consensus quality: 203603 bases at least Q40
Consensus quality: 206822 bases at least Q30
Consensus quality: 208592 bases at least Q20
Insert size: 214000; agarose-fp
Insert size: 211859; sum-of-contigs
Quality coverage: 5.34 in Q20 bases; agarose-fp
Quality coverage: 5.44 in Q20 bases; sum-of-contigs

* NOTE: This is a working draft sequence. It currently
* consists of 16 contigs. The true order of the pieces
* is not known and their order in this sequence record is
* arbitrary. Gaps between the contigs are represented as
* runs of N, but the exact sizes of the gaps are unknown.
* This record will be updated with the finished sequence
* as soon as it is available and the accession number will
* be preserved.

1 1400: contig of 1400 bp in length
1401 1500: gap of unknown length
1501 4621: contig of 3121 bp in length
4622 4721: gap of unknown length
4722 7894: contig of 3173 bp in length
7895 7994: gap of unknown length
7995 10806: contig of 2812 bp in length
10807 10906: gap of unknown length
10907 16818: contig of 5912 bp in length
16819 16918: gap of unknown length
16919 24662: contig of 7744 bp in length
24663 24762: gap of unknown length
24763 32389: contig of 7627 bp in length
32390 32490: gap of unknown length
32490 40414: contig of 7925 bp in length
40415 40514: gap of unknown length
40515 53087: contig of 12573 bp in length
53088 53187: gap of unknown length
53188 64676: contig of 11489 bp in length
64677 64776: gap of unknown length
64777 78536: contig of 13760 bp in length
78537 78636: gap of unknown length
78637 94647: contig of 16011 bp in length
94648 94747: gap of unknown length
94748 117462: contig of 22715 bp in length
117463 143519: contig of 25957 bp in length
143520 143619: gap of unknown length
143620 171176: contig of 27357 bp in length
171177 171276: gap of unknown length
171277 213359: contig of 42083 bp in length.

Location/Qualifiers
1..213359
/organism="Homo sapiens"

/db_xref="taxon:9606"
/chromosome="RP1-11"
/clone="RP11-773021"
1..1400
/note="assembly_name:Contig18"
1501..4621
/note="assembly_name:Contig19"
4722..7894
/note="assembly_name:Contig20"
7995..10806
/note="assembly_name:Contig21"
10907..16818
/note="assembly_name:Contig22"
16919..24662
/note="assembly_name:Contig23
clone_end:17
vector_side:right"
24763..32389
/note="assembly_name:Contig24"
32490..40414
/note="assembly_name:Contig25"
40515..53087
/note="assembly_name:Contig26"
53188..64676
/note="assembly_name:Contig27"
64777..78536
/note="assembly_name:Contig28"
78637..94647
/note="assembly_name:Contig29"
94748..117462
/note="assembly_name:Contig30"
117563..143519
/note="assembly_name:Contig31
clone_end:SP8
vector_side:left"
143620..171176
/note="assembly_name:Contig32"
171277..213359
/note="assembly_name:Contig33"

BASE COUNT 59337 a 47447 c 48595 g 56470 t 1506 others

ORIGIN

Query Match 100.0%; Score 210; DB 2; Length 213359;
Best Local Similarity 100.0%; Pred. No. 7.5e-54;
Matches 210; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 CCGCTTGTCTTCTCTCTCAAGCTGCAAGAAAGTCCAGAAAGCACAGCTGACTTAGG 60
|||||
DB 83670 CCGCTTGTCTTCTCTCTCAAGCTGCAAGAAAGTCCAGAAAGCACAGCTGACTTAGG 83611

QY 61 GAAGGCTGTGGAAATATCTCCGCTTTGGGGGGGCGAGGGGATGAGCGAGGGC 120
|||||
DB 83610 GAAGGCTGTGGAAATATCTCCGCTTTGGGGGGGCGAGGGGATGAGCGAGGGC 83551

QY 121 CGAGAAAGAACTGGAAGACTCCGTAGATTGCTTAGACCGCTCAGACACTCTCGGGCGC 180
|||||
DB 83550 CGAGAAAGAACTGGAAGACTCCGTAGATTGCTTAGACCGCTCAGACACTCTCGGGCGC 83491

QY 181 AGCGTGAGAGAGATTGTGCAACATTTC 210
|||||
DB 83490 AGCGTGAGAGATTGTGCAACATTTC 83461

RESULT 5
AX354807
LOCUS AX354807 2170 bp DNA linear PAT 06-FEB-2002
DEFINITION Sequence 1 from Patent WO0179290.
ACCESSION AX354807
VERSION AX354807.1 GI:18619538
KEYWORDS
SOURCE synthetic construct.
ORGANISM synthetic construct
artificial sequence.

REFERENCE 1 (sites)
 AUTHORS Drucker,D.J. and Lovshin,J.A.
 TITLE GLP-2 receptor gene promoter and uses thereof
 JOURNAL Patent: WO 017930-A 1 25-OCT-2001;
 1149336 ONTARIO INC. (CA)
 FEATURES Location/Qualifiers
 source 1. 2170
 /organism="synthetic construct"
 /db_xref="taxon:32630"
 /note="Recombinant DNA expression construct"
 BASE COUNT 574 a 500 c 551 g 540 t 5 others
 ORIGIN
 Query Match 29.6% Score 62.2; DB 6; Length 2170;
 Best Local Similarity 67.5%; Pred. No. 1.9e-08;
 Matches 141; Conservative 0; Mismatches 48; Indels 20; Gaps 3;
 QY 3 GCCTGTTCTTCTCTCCTGACCTGTCAAGAGAGTCCAGAAAGCACAGCTGACTT-AGGG 61
 1476 GTCTTCTTTCTTCTTCTGCGCTTCTGAGAGAGTCCAGGAGGAGTACGTTGGGG 1535
 Db 62 AAGCTCGGAAAAATCTCCCTGCTTTTGGGGGCGAGGGGCGGAGATAGCCAGGGCC 121
 1536 TAGGCTGGGAAAAATCTCCAGAGATTGTAGAGAGGGGCGAGGGGGGATGA----- 1586
 QY 122 GAGAAGAACTGTGAAGACTCCGATGATGCTCTAGACCGCTGACACTCTGGCGCA 181
 1587 -----GAAACTTGGAGATTCGGTAGATCGCTGTAGAGCAACTGACACAGTC-----GGCG 1636
 Db 182 GCCTGGAGAGATTGTGCAAACTTTCC 210
 QY 1637 GCCTGAAGAGACTTGTGCAAACTTCC 1665
 Db 1637 GCCTGAAGAGACTTGTGCAAACTTCC 1665
 RESULT 6 3130 bp DNA linear ROD 18-JUN-2001
 AF338224
 LOCUS Mus musculus glucagon-like peptide-2 receptor gene, partial cds.
 ACCESSION AF338224
 VERSION AF338224.1 GI:14485646
 KEYWORDS house mouse.
 SOURCE Mus musculus
 ORGANISM Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 Mammalia; Eutheria; Rodentia; Sclurognathi; Muridae; Murinae; Mus.
 1 (bases 1 to 3130)
 Lovshin,J.A., Estall,J., Yusta,B., Brown,T.J. and Drucker,D.J.
 Glucagon-like Peptide (GLP)-2 Action in the Murine Central Nervous
 System Is Enhanced by Elimination of GLP-1 Receptor Signaling
 J. Biol. Chem. 276 (24), 21489-21499 (2001)
 11262390
 JOURNAL 2 (bases 1 to 3130)
 PUBLISHED Lovshin,J.A. and Drucker,D.J.
 REFERENCE Direct Submission
 AUTHORS Submitted (17-JAN-2001) Department of Medicine, University of
 TITLE Toronto, 101 College Street CCRW3-845, Toronto, ON M5G2C4, Canada
 JOURNAL Location/Qualifiers
 source 1. 3130
 /organism="Mus musculus"
 /strain="129/SvJ"
 /db_xref="taxon:10090"
 2617. 2900
 /number=1
 2617. >2900
 /product="glucagon-like peptide-2 receptor"
 2835. >2900
 /codon_start=1
 /note="GLP-2 receptor"
 /product="glucagon-like peptide-2 receptor"
 /protein_id="AAK63043.1"
 /db_xref="GI:14485647"
 /translation="MRRLMGPTLPFLSLILVSIKO"
 BASE COUNT 805 a 741 c 769 g 797 t 18 others

ORIGIN
 Query Match 29.6% Score 62.2; DB 10; Length 3130;
 Best Local Similarity 67.5%; Pred. No. 1.9e-08;
 Matches 141; Conservative 0; Mismatches 48; Indels 20; Gaps 3;
 QY 3 GCCTGTTCTTCTCTCCTGACCTGTCAAGAGAGTCCAGAAAGCACAGCTGACTT-AGGG 61
 2436 GTCTTCTTTCTTCTTCTGCGCTTCTGAGAGAGTCCAGGAGGAGTACGTTGGGG 2495
 Db 62 AAGCTCGGAAAAATCTCCCTGCTTTTGGGGGCGAGGGGCGGAGATAGCCAGGGCC 121
 2496 TAGGCTGGGAAAAATCTCCAGAGATTGTAGAGAGGGGCGAGGGGGATGA----- 2546
 QY 122 GAGAAGAACTGTGAAGACTCCGATGATGCTCTAGACCGCTGACACTCTGGCGCA 181
 2547 -----GAAACTTGGAGATTCGGTAGATCGCTGTAGAGCAACTGACACAGTC-----GGCG 2596
 Db 182 GCCTGGAGAGATTGTGCAAACTTTCC 210
 2597 GCCTGAAGAGACTTGTGCAAACTTCC 2625
 Db 2597 GCCTGAAGAGACTTGTGCAAACTTCC 2625
 RESULT 7 207160 bp DNA linear HTG 01-MAR-2000
 AC016464
 LOCUS Mus musculus chromosome 11 clone RP23-409J21 map 11, WORKING DRAFT
 AC016464
 DEFINITION SFOURCE. 21 unordered pieces.
 AC016464.3 GI:7137319
 VERSION HTG: HTGS_PHASE1; HTGS_DRAFT.
 KEYWORDS house mouse.
 SOURCE Mus musculus
 ORGANISM Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 Mammalia; Eutheria; Rodentia; Sclurognathi; Muridae; Murinae; Mus.
 1 (bases 1 to 207160)
 Birren,B., Linton,L., Nusbaum,C. and Lander,E.
 JOURNAL Unpublished
 REFERENCE 2 (bases 1 to 207160)
 AUTHORS Birren,B., Linton,L., Nusbaum,C., Lander,E., Allen,N., Anderson,M.,
 Baldwin,J., Barna,N., Beckerly,R., Boguslavsky,L., Bouckhalter,B.,
 Brown,A., Castle,A., Colangelo,M., Collins,S., Collymore,A.,
 Cooke,P., DeRubeis,A., Dewar,K., Domino,M., Donelan,L., Doyle,M.,
 Ferreira,P., FitzHugh,W., Forrest,C., Funke,R., Gage,D.,
 Galagan,J., Gardyna,S., Grant,G., Hagos,B., Heaford,A., Horton,L.,
 Howland,J.C., Johnson,R., Jones,C., Kann,L., Karatas,A., Klein,J.,
 Lebeck,J., Lieu,C., Locke,K., Macdonald,P., Marquis,N.,
 McEwan,P., McGuck,A., McKernan,K., McLaughlin,J., Meldrum,J.,
 Morrow,J., Naylor,D., Norman,C.H., O'Connor,T., O'Donnell,P.,
 Peterson,K., Pollara,V., Riley,R., Roy,A., Santos,R., Severy,P.,
 Stange-Thomann,N., Stojanovic,N., Subramanian,A., Talamas,J.,
 Tesfaye,S., Tirrell,A., Vassiliev,H., Vo,A., Wheeler,J., Wu,X.,
 Wyman,D., Ye,W.J., Zimmer,A. and Zody,M.
 JOURNAL Direct Submission
 TITLE Submitted (30-NOV-1999) Whitehead Institute/MIT Center for Genome
 RESEARCH Research, 320 Charles Street, Cambridge, MA 02141, USA
 ON Mar 1, 2000 this sequence version replaced gi:6970326.
 All repeats were identified using RepeatMasker:
 Smit, A.F.A. & Green, P. (1996-1997)
 http://ftp.genome.washington.edu/RM/RepeatMasker.html
 ----- Genome Center
 Center: Whitehead Institute/ MIT Center for Genome Research
 Center code: WIDR
 Web site: http://www.seq.wi.mit.edu
 Contact: sequence.submissions@genome.wi.mit.edu
 ----- Project Information
 Center project name: L3198
 Center clone name: 409_J-21
 ----- Summary Statistics
 Sequencing vector: M13; M77815; 100% of reads
 Chemistry: Dye-terminator Big Dye; 100% of reads
 Assembly program: Phrap; version 0.960731

Consensus quality: 196848 bases at least Q40
 Consensus quality: 202204 bases at least Q30
 Consensus quality: 204082 bases at least Q20
 Insert size: 210000; agarose-fp
 Insert size: 205160; sum-of-ctrls
 Quality coverage: 4.8 in Q20 bases; agarose-fp
 Quality coverage: 4.9 in Q20 bases; sum-of-ctrls

NOTE: This is a 'working draft' sequence. It currently consists of 21 contigs. The true order of the pieces is not known and their order in this sequence record is arbitrary. Gaps between the contigs are represented as runs of N, but the exact sizes of the gaps are unknown. This record will be updated with the finished sequence as soon as it is available and the accession number will be preserved.

```
1
542 641: gap of 100 bp
642 1859: contig of 1218 bp in length
1860 1959: gap of 100 bp
1960 4152: contig of 2193 bp in length
4153 4252: gap of 100 bp
4253 6291: contig of 2039 bp in length
6292 6391: gap of 100 bp
6392 10537: contig of 4146 bp in length
10538 10637: gap of 100 bp
10638 14144: contig of 3507 bp in length
14145 14244: gap of 100 bp
14245 18479: contig of 4235 bp in length
18480 18579: gap of 100 bp
18580 25538: contig of 6959 bp in length
25539 25638: gap of 100 bp
25639 31995: contig of 6357 bp in length
31996 32095: gap of 100 bp
32096 40133: contig of 8038 bp in length
40134 40233: gap of 100 bp
40234 48264: contig of 8031 bp in length
48265 48364: gap of 100 bp
48365 55351: contig of 6987 bp in length
55352 55451: gap of 100 bp
55452 63846: contig of 8395 bp in length
63847 63946: gap of 100 bp
63947 75456: contig of 11510 bp in length
75457 75556: gap of 100 bp
75557 88148: contig of 12592 bp in length
88149 88248: gap of 100 bp
88249 99481: contig of 11233 bp in length
99482 99581: gap of 100 bp
99582 115973: contig of 16392 bp in length
115974 116073: gap of 100 bp
116074 132294: contig of 16221 bp in length
132295 13394: gap of 100 bp
13395 149262: contig of 16868 bp in length
149263 149362: gap of 100 bp
149363 177577: contig of 28215 bp in length
177578 177677: gap of 100 bp
177678 207160: contig of 29483 bp in length.
```

FEATURES

source

```
1..207160
/organism="Mus musculus"
/db_xref="taxon:10090"
/chromosome="11"
/map="11"
/clone="RP23-409J21"
/clone_lib="RPC1-23 Female Mouse BAC"
1..541
/note="assembly-fragment"
clone_end:17
vector_side:right"
misc_feature
642..1859
/note="assembly-fragment"
misc_feature
1960..4152
/note="assembly-fragment"
```

```
misc_feature 4253..6291
/note="assembly-fragment"
misc_feature 6392..10537
/note="assembly-fragment"
misc_feature 10638..14144
/note="assembly-fragment"
misc_feature 14245..18479
/note="assembly-fragment"
misc_feature 18580..25538
/note="assembly-fragment"
misc_feature 25639..31995
/note="assembly-fragment"
misc_feature 32096..40133
/note="assembly-fragment"
misc_feature 40234..48264
/note="assembly-fragment"
misc_feature 48365..55351
/note="assembly-fragment"
misc_feature 55452..63846
/note="assembly-fragment"
misc_feature 63947..75456
/note="assembly-fragment"
misc_feature 75557..88148
/note="assembly-fragment"
misc_feature 88249..99481
/note="assembly-fragment"
misc_feature 99582..115973
/note="assembly-fragment"
misc_feature 116074..132294
/note="assembly-fragment"
misc_feature 132395..149262
/note="assembly-fragment"
misc_feature 149363..177577
vector_side:right"
misc_feature 177678..207160
/note="assembly-fragment"
misc_feature 207160
/note="assembly-fragment"
BASE COUNT 57689 a 46293 c 45874 g 55300 t 2004 others
ORIGIN
```

Query Match 29.6%; Score 62.2; DB 2; Length 207160;
 Best Local Similarity 67.5%; Pred.No.1.9e-08;
 Matches 141; Conservative 0; Mismatches 48; Indels 20; Gaps 3;

```
QY 3 GCCTGTTCTTCTCTCTCAGCCGTCAGCAATCCCAAGACACAGCTGACTT-AGGG 61
DB 28673 GCTTGCCTTTCTTCTTGGCTTGGAGAGAGTCCAGGACAGCTAGAGCTTGGGGG 28732
QY 62 AAGCTCTGGAAATATCTCCCTGCTTTGGGGGCGCAGGGGCGGGGATGAGCGACGGCC 121
DB 28733 TAGGCTCTGGAAATATCTCCCAAGATTTAGAGAGGCGCAGCGGGGATGA----- 28783
QY 122 GAGAAAGAACTCTGAAGACTCCGTAGATTGCTAGACCGCCTCAGACACTCTCGGCGCA 181
DB 28784 -----GAAACTTGGAGATTGCGTAGATCGCTGTAGAGCAACTCAGACAGTTC---GGCG 28833
QY 182 GCGTGAAGAGATTTGTGCAACATTTCC 210
DB 28834 GCTGAAGAGACTTGTGCAACACTTCC 28862
```

```
RESULT 8
AL646097 211173 bp DNA linear HTG 10-JAN-2002
LOCUS AL646097 Mus musculus chromosome 11 clone RP23-338M9, *** SEQUENCING IN
DEFINITION PROGRESS ***, in unordered pieces.
ACCESSION AL646097
VERSION AL646097.5 GI:18135259
KEYWORDS HTG: HTGS_PHASE1; HTGS_DRAFT.
SOURCE house mouse.
ORGANISM Mus musculus
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
```

REFERENCE	Mammalia: Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus
AUTHORS	McLay,K.
TITLE	Direct Submission
JOURNAL	Submitted (08-JAN-2002) Wellcome Trust Sanger Institute, Hinxton, Cambridgeshire, CB10 1SA, UK. E-mail enquiries: humquery@sanger.ac.uk Clone requests: clonerequest@sanger.ac.uk On Jan 11, 2002 this sequence version replaced gi118072579.
COMMENT	----- Center: Wellcome Trust Sanger Institute Genome Center ----- Center code: SC Web site: http://www.sanger.ac.uk Contact: humquery@sanger.ac.uk ----- Project Information ----- Center project name: BM338M9 ----- Summary Statistics ----- Assembly program: XGAP4; version 4.5 Sequencing vector: plasmid; L08752; 100% of reads Chemistry: Dye-terminator Big Dye; 100% of reads Consensus quality: 206626 bases at least Q40 Consensus quality: 207698 bases at least Q30 Consensus quality: 208567 bases at least Q20 Insert size: 209573; sum-of-contigs Insert size: 207120; 2.8% error; agarose-fp Quality coverage: 9.05x in Q20 bases; sum-of-contigs quality coverage: 9.26x in Q20 bases; agarose-fp ----- * NOTE: This is a 'working draft' sequence. * This record will be updated with the finished sequence * as soon as it is available and the accession number will * be preserved. -----
FEATURES	Location/Qualifiers 1..211173 /organism="Mus musculus" /db_xref="taxon:10090" /chromosome="11" /clone="RP23-338M9" /clone_lib="RPCI-23" 1..8490 /note="assembly_fragment:04365 fragment_chain:1" 8591..10949 /note="assembly_fragment:02882 fragment_chain:1" 11050..22821 /note="assembly_fragment:01376 fragment_chain:1" 22922..39369 /note="assembly_fragment:01658 fragment_chain:1" 39470..47112 /note="assembly_fragment:00263 fragment_chain:1" 47213..51333 /note="assembly_fragment:02994 fragment_chain:2" 51434..56653 /note="assembly_fragment:00672 fragment_chain:2" 56754..71877 /note="assembly_fragment:03096 fragment_chain:2" 71978..112160 /note="assembly_fragment:01861 fragment_chain:2" 11261..129027 /note="assembly_fragment:01143 fragment_chain:2" 129128..142170 /note="assembly_fragment:00621 fragment_chain:2" 142271..146192 /note="assembly_fragment:00072

	misc_feature	fragment_chain:2" 146293..185554 /note="assembly_fragment:01439 fragment_chain:2"
	misc_feature	185655..187770 /note="assembly_fragment:00832 fragment_chain:2"
	misc_feature	187871..195431 /note="assembly_fragment:02362 fragment_chain:2"
	misc_feature	195532..200953 /note="assembly_fragment:03683 fragment_chain:2"
	misc_feature	201054..211173 /note="assembly_fragment:01683 fragment_chain:2 clone_end:T7 vector_side:right"
BASE COUNT	57360 a 48310 c 47762 g 56133 t 1608 others	
ORIGIN		
Query Match	29.6%; Score 62.2; DB 2; Length 211173;	
Best Local Similarity	67.5%; Pred. No. 1.9e-08;	Mismatches 48; Indels 20; Gaps 3;
Matches 141:	Conservative 0;	
Oy	3 GCCTGTCCTTTCACCCAGCCTGTGCAGAAGATGCCAAGAACACACTGACTT-AAGG 61	
Dd	132359 GCCTTGGCTTTTTCTCTCGTGCAAGCATGGAAGTCCCAGCACAGCGTAGAGCTTTGGGG 132300	
Oy	62 AAGGCTGTGGAAAAATTCTCCCTGCTTTTGGSGGGGACAAGCGCGGGGCGATGACCAGGGCC 121	
Dd	132299 TAAAGCTGGGAAAATCTCCCCAAGATTATTAGAGGGGCGAGCGGGGATGA----- 132249	
Oy	122 GAAGAAGAACTGGAAGACTCGGTGATTGGCTCTAGACCGCTCAGACACTCTGCGCCA 181	
Dd	132248 -----GAAACTGTGAGATTCGTGATTCGCTGTAGACCACTCAGACATC----GCGC 132199	
Oy	182 GCGTGGAGAGATTGTGCAAAACATTTC C 210	
Dd	132198 GCCTGAAGAGACTTGTCAAACACTTCC 132170	
RESULT 9		
AX354812	AX354812	493 bp DNA linear PAT 06-FEB-2002
LOCUS	Sequence 6 from Patent WO0179290.	
DEFINITION	AX354812	
ACCESSION	AX354812	
VERSION	AX354812.1 GI:18619543	
KEYWORDS	"Mus sp."	
SOURCE	Mus sp.	
ORGANISM	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.	
REFERENCE	Drucker,D.J. and Lovshin,J.A. 1 (sites)	
AUTHORS	GIP-2 receptor gene promoter and uses thereof	
TITLE	Patent: WO 0179290-A 6 25-OCT-2001;	
JOURNAL	NAR INC. (CA) 1149336 ONPARIO Inc.	
FEATURES	Location/Qualifiers	
source	1..493	
CDS	/organism="Mus sp." /db_xref="taxon:10095" 401.. /note="unnamed protein product"	
	/codon_start=1 /protein_id="CAD22992.1" /db_xref="GI:18619544" /translation="MRRLMGPGPFSLILLVTSIKQ"	
BASE COUNT	100 a 132 c 155 g 104 t 2 others	
ORIGIN		
Query Match	25.7%; Score 54; DB 6; Length 493;	
Best Local Similarity	62.0%; Pred. No. 6.3e-06;	

```

Matches 129: Conservative 0; Mismatches 60; Indels 19; Gaps 2;

OY 3 GCCTGTTCTTCTCCAGCCTGTCAGAGATCCAGAAAGACAGCTGACTTAAAGGA 62
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 3 GTCTTGCTTTTCTTCTGCGCTTCTAGAGAGTCCAGAGCAGTACGCTCTGGGGG 62
OY 63 AGCTCTGGAAAAATCTCCCTGCTTTTGGGGGGGAGAGGGGATGAGCCAGGGCCG 122
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 63 TAGCTCTGGAAAAATCTCCAGAGATTTAGAGAGGGGAGGGGGATGA----- 112
OY 123 AGAAGAACTCTAGAGACCTGCTAGAGACCGCTTACAGACTCTGGCGGAG 182
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 113 -----GAAACTTGAGATTGCTAGATCGCTGTAAGAGCAACACAGCTCG-----GCCG 163
OY 183 CGTGAGAGAGCATTTGTGCAAAACATTTCC 210
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 164 CCGAAGAGGACTTTGTGCAAAACACTTCC 191

RESULT 10
AK021842/c
LOCUS
DEFINITION Homo sapiens CDNA FLJ11780 fls, clone HEMBA1005931, weakly similar
to ZINC FINGER PROTEIN 83.
ACCESSION AK021842
VERSION AK021842.1 GI:10433110
KEYWORDS oligo capping, fls (full insert sequence).
SOURCE Homo sapiens embryo, 10 weeks whole embryo, mainly head CDNA to
Homo, clone.lib:HEMBA1 clone:HEMBA1005931.
ORGANISM Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
REFERENCE
AUTHORS Nishikawa,T., Ota,T., Hayashi,K., Sugiyama,T., Otsuki,T., Suzuki,Y.,
Magatsuna,M., Hosokiri,T., Kaku,Y., Kodaira,H., Kondo,H.,
Sugawara,M., Takahashi,M., Chiba,Y., Ishida,S., Murakawa,K.,
Ono,Y., Takiguchi,S., Watanabe,S., Kimura,K., Murakami,K.,
Ishii,S., Kawai,Y., Saito,K., Yamamoto,J., Makamatsu,A.,
Nakamura,Y., Nagahara,K., Masuno,Y., Niinomiya,K. and Iwayanagi,T.
NEDO human CDNA sequencing project
TITLE Unpublished (2000)
JOURNAL 2 (phases 1 to 1612)
AUTHORS Isogai,T. and Otsuki,T.
REFERENCE Direct Submission
JOURNAL Submitted (23-AUG-2000) to the DDBJ/EMBL/GenBank databases. Takao
Isogai, Helix Research Institute, Genomics Laboratory, 1532-3 Yana,
Kisarazu, Chiba 292-0812, Japan (E-mail: genomics@hri.co.jp,
Tel:81-438-52-3951, Fax:81-438-52-3952)
COMMENT International Trade and Industry of Japan: CDNA full insert
sequencing: Research Association for Biotechnology: CDNA library
construction, 5'- & 3'-end one pass sequencing and clone selection:
Helix Research Institute (supported by Japan Key Technology Center
etc.) and Department of Virology, Institute of Medical Science,
University of Tokyo.
FEATURES
source
Location/Qualifiers
1..1612
/organism="Homo sapiens"
/db_xref="taxon:9606"
/clone="HEMBA1005931"
/dev_stage="embryo, 10 weeks"
/tissue_type="whole embryo, mainly head"
/note="cloning vector: pME18SFL3"

BASE COUNT 352 a 476 c 391 g 393 t

Query Match 18.4%; Score 38.6; DB 9; Length 1612;
Best Local Similarity 52.1%; Pred. NO. 0.34;
Matches 86; Conservative 0; Mismatches 79; Indels 0; Gaps 0;

37. CCCAGAAAGCAGCTAGTAGGAGAGCTCGGAAATATCTCCGCTTTGGGGGG 96
misc-feature
1..53967
/organism="Homo sapiens"
/db_xref="taxon:9606"
/chromosome="1"
/clone="RP11-415J8"
/clone.lib="RPC1-11.2"
1..53967
/note="assembly fragment:03338
fragment_chain:1"

```

```

Db 258 CCCAAGACACAGAGGACCACTGGGAAAGACAGGGAGACGTCCTCCACCTGGGACCCCT 199
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
OY 97 CAGGGGCGGGGATGAGCCAGGCGGAGAGAGATCTGAAGATCGGTAGATTGCTTA 156
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 198 TAAATCTGGAGAGAGAGAGCTTTCCTGCGGGGAGACATTACAGAGCTTGACATAGCTTA 139
OY 157 GACCGGCTCAGACATCTCCGCGGAGCGCTGAGAGAGATTGTGCA 201
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 138 CACCTCATAGGAGAGAGAGCGCTGATCTCGGTGAGGGTTTCGCCA 94

RESULT 11
AL513327/c
LOCUS
DEFINITION Homo sapiens chromosome 1 clone RP11-415J8, *** SEQUENCING IN
PROGRESS ***; 3 unordered pieces.
ACCESSION AL513327
VERSION AL513327.22 GI:18476646
KEYWORDS HTG; HTGS_PHASE1; HTGS_ACTIVEFIN; HTGS_DRAFT; HTGS_FULLTOP.
SOURCE human.
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
REFERENCE
AUTHORS Matthews,N.
TITLE Direct Submission
JOURNAL Submitted (24-JAN-2002) Wellcome Trust Sanger Institute, Hinxton,
Cambridgeshire, CB10 1SA, UK. E-mail enquiries:
humquery@sanger.ac.uk Clone requests: clonequest@sanger.ac.uk
On Feb 1, 2002 this sequence version replaced g1:17384473.
COMMENT ----- Genome Center
Center: Wellcome Trust Sanger Institute
Center code: SC
Web site: http://www.sanger.ac.uk
Contact: humquery@sanger.ac.uk
----- Project Information
Center project name: BA41508
----- Summary Statistics
Assembly program: XGAP4; version 4.5
Sequencing vector: plasmid; L08752; 100% of reads
Chemistry: dye-terminator Big Dye; 100% of reads
Consensus quality: 202251 bases at least Q40
Consensus quality: 202335 bases at least Q30
Insert size: 202482; sum-of-contigs
Insert size: 196528; 5.3% error; agarose-fp
Quality coverage: 10.24x in Q20 bases; sum-of-contigs Quality
coverage: 10.51x in Q20 bases; agarose-fp
-----
* NOTE: This is a 'working draft' sequence. It currently
* consists of 3 contigs. The true order of the pieces
* is not known and their order in this sequence record is
* arbitrary. Gaps between the contigs are represented as
* runs of N, but the exact sizes of the gaps are unknown.
* This record will be updated with the finished sequence
* as soon as it is available and the accession number will
* be preserved.
1..53967: contig of 53967 bp in length
* 53968 54067: gap of 100 bp
* 54068 147217: contig of 93150 bp in length
* 147218 147317: gap of 100 bp
* 147318 202682: contig of 55365 bp in length.
Location/Qualifiers
1..202682
/organism="Homo sapiens"
/db_xref="taxon:9606"
/chromosome="1"
/clone="RP11-415J8"
/clone.lib="RPC1-11.2"
1..53967
/note="assembly fragment:03338
fragment_chain:1"

```

```

misc_feature      54068..147217
                  /note="assembly_fragment:06038
                  fragment_chain:1"
misc_feature      147318..202682
                  /note="assembly_fragment:05678
                  fragment_chain:1"

BASE COUNT      52289 a 47427 c 48564 g 54202 t      200 others
ORIGIN

Query Match      18.4%: Score 38.6; DB 2; Length 202682;
Best Local Similarity 52.1%: Pred. No. 0.33;
Matches 86; Conservative 0; Mismatches 79; Indels 0; Gaps 0;

QY 37 CCCAGAACACACAGCTGACTAGTACAGAGCTCTGTGAAAAATCTCCCTGCTTTGGGGGGG 96
    ||| ||||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB 78763 CCCAGAACACACAGAGCAGCAGCTGAGAAAGACAGGAGACATCTCCACGCTGGGACACCT 78704
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 97 CAGGGCGGGGGGATGAGCGAGGCGCGAAGAACTCTGTAAGACTCCGTAGATTGCTCTA 156
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB 78703 TAAATCTGAGAGGAGAACTCTGCTGGGGGAAACATTACAGAGCTGAGAGATAGGTTA 78644
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 157 GACCGCTCAGACACTCTCTGCGCGCAGCTGAGAGAGATTGTGCA 201
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB 78643 CACCTCATAGGAGAAAGAGCGCTGATCTGTGTGAGGTTGCCCA 78599
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||

RESULT 12
AC106475      147868 bp      DNA      linear      HMG 12-JAN-2002
LOCUS      Rattus norvegicus clone CH230-160H18, *** SEQUENCING IN PROGRESS
DEFINITION      *** 75 unordered pieces.
ACCESSION      AC106475
VERSION      AC106475.1 GI:18138997
KEYWORDS      HMG; HMG5-PHASE1.
SOURCE      Norway rat.
ORGANISM      Rattus norvegicus
              Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
              Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae;
              Rattus.
REFERENCE      1 (bases 1 to 147868)
AUTHORS      Muzny,D.M., Adams,C., Adio-Oduola,B., All-osman,F.R., Allen,C.,
              Alstrooms,S.L., Amaratunga,H.C., Are,J.R., Banks,T., Barberia,J.,
              Benton,J., Blincoe,K., Blankenburg,K., Bonnin,D., Bouck,J.,
              Bowie,S., Brieva,M., Brown,E., Brown,M., Bryant,N.P., Buhay,C.,
              Burch,P., Burkett,C., Butrell,K.L., Byrd,N.C., Caron,T.F.,
              Carter,M., Cavazos,S.R., Chacko,J., Chavez,D., Chen,G., Chen,R.,
              Chen,Z., Chowdhury,I., Christopoulos,C., Cleveland,C.D., Cox,C.,
              Coyle,M.D., Dathorne,S.R., David,R., Davila,M.L., Davis,C.,
              Denny,A.L., Ding,Y., Dinh,H.H., Douthwaite,K.J., Draper,H.,
              Dugan-Rocha,S., Durbin,K.J., Earnhart,C., Edgar,D., Edwards,C.C.,
              Elhaj,C., Escotto,M., Falls,T., Ferraguto,D., Flagg,N., Ford,J.,
              Foster,P., Frantz,P., Gabisi,A., Gao,J., Garcia,A., Garner,T.,
              Garza,N., Gill,R., Gorrell,J.H., Guevara,W., Gunaratne,P., Hale,S.,
              Hamilton,K., Harris,C., Harris,K., Hart,M., Havlak,P., Hawes,A.,
              Hernandez,J., Hernandez,O., Hodgson,A., Hognes,M., Holloway,C.,
              Hollins,B., Homs,F., Howard,S., Huber,J., Huliy,S., Hume,J.,
              Jackson,L.E., Jacobson,B., Jia,Y., Johnson,R., Jolivet,S.,
              Jonathan,S., Karlsson,E., Kelly,S., Khan,U., King,L., Korvah,J.,
              Kovar,C., Kratovic,J., Kureshi,A., Landry,N., Leal,B., Lewis,L.C.,
              Lewis,L., Li,U., Li,Z., Lichtarge,O., Lieu,C., Liu,J., Liu,W.,
              Louisedge,H., Lozado,R.J., Lu,X., Lucier,A., Lucier,R., Luna,R.,
              Ma,J., Maheshwari,M., Mapua,P., Martin,R., Martindale,A.,
              Martinez,E., Massey,E., Mawhinney,E., McLeod,M.P., Meador,M.,
              Mel,G., Metzker,M., Miner,G., Miner,Z., Mitchell,T., Mohabdat,K.,
              Morgan,M., Morris,S., Moser,M., Neal,D., Newton,J., Newton,N.,
              Nguyen,A., Nguyen,N., Nguyen,N., Nickerson,E., Nwokwenko,S.,
              Oguh,M., Okunou,G., Oragunye,N., Oviedo,R., Pace,A., Payton,B.,
              Peery,J., Perez,L., Peters,L., Pichens,R., Primus,E., Pu,L.L.,
              Quiles,M., Ren,Y., Rives,M., Rojas,A., Rojudoan,I., Rolfe,M.,
              Ruiz,S., Savery,G., Scherer,S., Scott,G., Shen,H., Shoshitaishvili,N.,
              Sisson,I., Sodergren,E., Sonalke,T., Sparks,A., Stanley,H.,
              Stone,H., Sutton,A., Svatek,A., Tabor,P., Tamerisa,A., Tamerisa,K.,

```

```

Tang,H., Tansey,J., Taylor,C., Taylor,T., Telford,B., Thomas,N.,
Thomas,S., Usmani,K., Vasquez,L., Vera,V., Villalón,D., Vinson,R.,
Wall,R., Wang,S., Ward-Moore,S., Warren,R., Washington,C.,
Watlington,S., Williams,G., Williamson,A., Wleciyk,R., Wooden,S.,
Worley,K., Wu,C., Wu,Y., Zhou,J., Zorrilla,S., Nelson,D.,
Weinstock,G. and Gibbs,R.

Direct Submission
Unpublished
2 (bases 1 to 147868)
Worley,K.C.
Center: Baylor College of Medicine
Center code: BCM
Web site: http://www.hgsc.bcm.tmc.edu/
Contact: hgsc-help@bcm.tmc.edu
----- Project Information
Center project name: GK2D
Center clone name: CH230-160H18
----- Summary Statistics
findPhrapList
Assembly program: Phrap; version 0.990329First call to
Consensus quality: 102886 bases at least Q40
Consensus quality: 113109 bases at least Q30
Consensus quality: 121647 bases at least Q20
Estimated insert size: 94980; sum-of-configs estimation
Quality coverage: 0x in Q20 bases; agatose-fp estimation
Quality coverage: 1.1x in Q20 bases; sum-of-configs estimation
-----
* NOTE: Estimated insert size may differ from sequence length
* (see http://www.hgsc.bcm.tmc.edu/docs/genbank/draft_data.html).
* NOTE: This is a 'working draft' sequence. It currently
* consists of 75 contigs. The true order of the pieces
* is not known and their order in this sequence record is
* arbitrary. Gaps between the contigs are represented as
* runs of N, but the exact sizes of the gaps are unknown.
* This record will be updated with the finished sequence
* as soon as it is available and the accession number will
* be preserved.
1
8024      8023: contig of 8023 bp in length
8124      8123: gap of unknown length
11675     11675: contig of 3552 bp in length
11776     11775: gap of unknown length
15149     15149: contig of 3374 bp in length
15249     15249: gap of unknown length
15250     15250: contig of 3925 bp in length
15175     15175: gap of unknown length
19275     19274: gap of 3368 bp in length
22643     22742: gap of unknown length
22743     22742: gap of 3106 bp in length
25848     25848: gap of unknown length
25948     25948: gap of unknown length
28042     28042: contig of 2094 bp in length
28043     28042: gap of unknown length
29999     29999: contig of 1857 bp in length
30000     30000: gap of unknown length
30100     30099: gap of unknown length
33397     33397: contig of 3298 bp in length
33398     33397: gap of unknown length
33498     33497: gap of unknown length
36954     36953: contig of 3456 bp in length
37054     37053: gap of unknown length
39928     39928: contig of 2875 bp in length
40028     40028: gap of unknown length
40029     40028: gap of 2108 bp in length
42136     42136: contig of 2108 bp in length
42137     42136: gap of unknown length
42237     42236: contig of 3330 bp in length
45566     45566: gap of unknown length
45667     45667: gap of 1487 bp in length
47154     47153: gap of unknown length
47254     47253: contig of 1645 bp in length
48898     48898: contig of 1645 bp in length
48899     48898: gap of unknown length
48999     48998: gap of 1302 bp in length
50300     50300: contig of 1302 bp in length

```

```

50301 50400: gap of unknown length
50401 52989: contig of 2589 bp in length
52990 54089: gap of unknown length
54090 54605: contig of 1516 bp in length
54606 54705: gap of unknown length
54706 57531: contig of 2826 bp in length
57532 60066: contig of 2435 bp in length
60067 60167: gap of unknown length
60168 61917: contig of 1651 bp in length
61918 63469: contig of 1552 bp in length
63470 65866: gap of unknown length
65867 65966: contig of 2297 bp in length
65967 68814: contig of 2848 bp in length
68815 70324: gap of unknown length
70325 70424: gap of unknown length
70425 72206: contig of 1782 bp in length
72207 72306: gap of unknown length
72307 73955: contig of 1649 bp in length
73956 74055: gap of unknown length
74056 75353: contig of 1298 bp in length
75354 75453: gap of unknown length
75454 77676: contig of 2223 bp in length
77677 79896: contig of 2120 bp in length
79897 81516: gap of unknown length
81517 81616: gap of unknown length
81617 83235: contig of 1619 bp in length
83236 85103: contig of 1768 bp in length
85104 86799: gap of unknown length
86799 86899: contig of 1596 bp in length
86900 88102: gap of unknown length
88103 89377: contig of 1175 bp in length
89378 90752: gap of unknown length
90753 90852: gap of unknown length
90853 92238: contig of 1386 bp in length
92239 92339: gap of unknown length
92340 94149: contig of 1811 bp in length
94150 94249: gap of unknown length
94250 96535: contig of 2286 bp in length
96536 96635: gap of unknown length
96636 98476: contig of 1841 bp in length
98477 99663: contig of 1087 bp in length
99664 101115: gap of unknown length
101116 101215: contig of 1352 bp in length
101216 102350: gap of unknown length
102351 102630: gap of unknown length
102631 104194: contig of 1564 bp in length
104195 104294: gap of unknown length
104295 105839: contig of 1545 bp in length
105840 105939: gap of unknown length
105940 107145: contig of 1206 bp in length
107146 107245: gap of unknown length
107246 108445: contig of 1200 bp in length
108446 110126: contig of 1581 bp in length
110127 110226: gap of unknown length
110227 111821: contig of 1595 bp in length
111822 111921: gap of unknown length
111922 113705: contig of 1784 bp in length
113706 113805: gap of unknown length
113806 115339: contig of 1534 bp in length
115340 115439: gap of unknown length

```

```

Query Match      17.5%; Score 36.8; DB 2; Length 147868;
Best Local Similarity 59.6%; Pred. No. 1.2;
Matches 62; Conservative 0; Mismatches 42; Indels 0; Gaps 0;

QY 92 GGGGCGAGGGGGGGGGGATGAGCCAGGCGCGAGAGAACTCTGAACACTCCGATTTG 151
      ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 144736 GGGGCGAGGGGGGGGGGATGAGCCAGGCGCGAGAGAACTCTGAACAGGGCGGGGTTG 144795
      ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||

QY 152 CTTTACAGCCCTCAGACACTCTGCGGCGAGCGTGGAGAGATT 195
      ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 144796 CCGGAGAGGGGGCTAATTAACCTTGCGATAGCGGGGGGATT 144839
      ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||

```

```

RESULT 13
AC096681/c
LOCUS
DEFINITION
AC096681
Canis familiaris clone RP81-228C19, WORKING DRAFT SEQUENCE, 3
AC096681
VERSION
AC096681.1 GI:15718555
HTG: HTGS_PHASE1; HTGS_DRAFT.
KEYWORDS
SOURCE
dog.
ORGANISM
Canis familiaris
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Euteria; Carnivora; Fissipedia; Canidae; Canis.
1 (bases 1 to 145617)
Akhter,N., Ayele,K., Beckstrom-Stenbery,S.M., Benjamin,B.,
Blakesley,R.W., Bouffard,G.G., Green,K., Brinkley,C., Brooks,S.,
Dietrich,N.L., Grant,S., Guan,X., Gupta,J., Haghighi,P.,
Ho,S.-L., Ido,J.R., Karlins,E., Lario,P., Lee-Lin,S.-O.,
Legaspi,R., Maduro,Q.L., Maduro,V.B., Mastello,C., Mastrian,S.D.,
Mcloskey,J.C., McDowell,J., Pearson,R., Prasad,A., Shevchenko,Y.,
Stattinop,S., Thomas,J.W., Thomas,P.J., Touchman,D.W.,
Tsurguev,C., Vogt,J.L., Walker,M.A., Wetherby,K.D., Wiggins,L.,
Zhang,L.-H. and Green,E.D.
NISC Comparative Sequencing Initiative
Unpublished
2 (bases 1 to 145617)
Green,E.D.
Direct Submission
Submitted (21-SEP-2001) NIH Intramural Sequencing Center, 8717
Groveomont Circle, Gaithersburg, MD 20877, USA
Genome Center

```

```

TITLE
JOURNAL
REFERENCE
AUTHORS
TITLE
JOURNAL
COMMENT
Center: NIH Intramural Sequencing Center
Center code: NISC
Web site: http://www.nisc.nih.gov
Contact: nisc.mouse@nigr.nih.gov
----- Project Information
Center project name: cti
Center clone name: 228C19
----- Summary Statistics
Sequencing vector: plasmid; n/a; 100% of reads
Chemistry: Dye-terminator Big Dye; 100% of reads
Assembly program: Phrap; version 0.990319
Consensus quality: 144648 bases at least Q40
Consensus quality: 144809 bases at least Q30
Consensus quality: 144939 bases at least Q20
Insert size: 145000; agarose-fp
Insert size: 145417; sum-of-ctdigs
Quality coverage: 10.21x in Q20 bases; agarose-fp
Quality coverage: 10.18x in Q20 bases; sum-of-ctdigs
-----

```

* NOTE: This is a 'working draft' sequence. It currently

----- Genome Center -----
 Center: Washington University Genome Sequencing Center
 Center code: WUGSC
 Web site: <http://genome.wustl.edu/gsc/index.shtml>
 ----- Project Information -----

[illegible]

GenCore version 5.1.3
Copyright (c) 1993 - 2002 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: November 6, 2002, 14:05:40 ; Search time 7.97326 Seconds
(without alignments)
6469.502 Million cell updates/sec

Title: US-09-833-740-7_COPY_1_210
Perfect score: 210
Sequence: 1 ccgcctgtcttcttcttccccc.....ggattgtgcaaacatttcc 210

Scoring table: IDENTITY_NUC
Gapop 10.0 , Gapext 1.0

Searched: 38353 seqs, 122816752 residues

Total number of hits satisfying chosen parameters: 767066

Minimum DB seq length: 0
Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%
Maximum Match 100%

Listing first 45 summaries

Database : Issued Patents, NA: *
1: /cgn2_6/ptodata/1/ina/5A.COMB.seq:*
2: /cgn2_6/ptodata/1/ina/5B.COMB.seq:*
3: /cgn2_6/ptodata/1/ina/6A.COMB.seq:*
4: /cgn2_6/ptodata/1/ina/6B.COMB.seq:*
5: /cgn2_6/ptodata/1/ina/PCRNUS.COMB.seq:*
6: /cgn2_6/ptodata/1/ina/backfile1.seq:*

Pred. No. is the number of results predicted by chance to have a
score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	29	13.8	2674	4	US-09-817-180-1
2	28.4	13.5	8791	5	PCT-US96-01735-5
3	28	13.3	4791	4	US-08-949-155-49
4	27.8	13.2	3228	4	US-09-318-448-20
5	27.8	13.2	4184	2	US-08-785-310A-4
6	27	12.9	1762	6	5185262-2
7	27	12.9	1763	6	5185262-1
8	26.8	12.8	3740	4	US-09-162-274A-6
9	26.6	12.7	7218	1	US-08-232-463-14
10	26.4	12.6	584	4	US-09-328-111-83
11	26.4	12.6	1489	4	US-09-487-445-10
12	26.2	12.5	7210	2	US-08-257-963B-10
13	26.2	12.5	7210	4	US-08-367-841A-10
14	26.2	12.5	7210	5	PCT-US95-07201-10
15	26.2	12.5	22481	4	US-08-367-841A-43
16	26.2	12.5	22481	5	PCT-US95-07201-43
17	26.2	12.5	35060	3	US-08-814-095-7
18	26	12.4	26	3	US-08-845-54C-23
19	26	12.4	1392	4	US-09-006-353A-1
20	26	12.4	1933	4	US-08-974-380-1
21	26	12.4	1968	5	PCT-US95-14024-2
22	26	12.4	2153	4	US-08-577-492-31
23	26	12.4	2153	4	US-09-079-630-31
24	26	12.4	6407	2	US-08-616-844-7
25	26	12.4	6407	2	US-08-599-654-7
26	26	12.4	6407	2	US-08-944-868A-7
27	26	12.4	6407	3	US-08-944-423A-7

28	26	12.4	6407	3	US-08-944-496-7	Sequence 7, Appl
29	26	12.4	8106	4	US-09-135-241-1	Sequence 1, Appl
30	25.8	12.3	972	1	US-07-940-605A-9	Sequence 9, Appl
31	25.8	12.3	972	2	US-08-690-096-9	Sequence 8, Appl
32	25.8	12.3	2301	4	US-09-085-199B-8	Sequence 6, Appl
33	25.6	12.2	548	4	US-08-469-667-6	Sequence 6, Appl
34	25.6	12.2	548	4	US-09-224-110-6	Sequence 6, Appl
35	25.6	12.2	548	5	PCT-US95-07289-6	Sequence 6, Appl
36	25.6	12.2	1067	4	US-09-045-193-1	Sequence 1, Appl
37	25.6	12.2	2517	1	US-07-906-930E-1	Sequence 1, Appl
38	25.6	12.2	2939	1	US-07-906-930E-3	Sequence 3, Appl
39	25.6	12.2	15297	4	US-09-817-180-3	Sequence 3, Appl
40	25.6	12.2	51952	3	US-08-947-823-1	Sequence 1, Appl
41	25.4	12.1	434	1	US-08-480-784-34	Sequence 34, Appl
42	25.4	12.1	434	1	US-08-483-553-34	Sequence 34, Appl
43	25.4	12.1	434	1	US-08-487-002-34	Sequence 34, Appl
44	25.4	12.1	434	1	US-08-483-554B-34	Sequence 34, Appl
45	25.4	12.1	434	1	US-08-488-011B-34	Sequence 34, Appl

ALIGNMENTS

RESULT 1
US-09-817-180-1
; Sequence 1, Application US/09817180
; Patent No. 6340584
; GENERAL INFORMATION:
; APPLICANT: GAN, Weiniu et al.
; TITLE OF INVENTION: ISOLATED HUMAN KINASE PROTEINS, NUCLEIC
; TITLE OF INVENTION: ACID MOLECULES ENCODING HUMAN KINASE PROTEINS, AND USES
; FILE REFERENCE: CLO001183
; CURRENT APPLICATION NUMBER: US/09/817,180
; CURRENT FILING DATE: 2001-03-27
; NUMBER OF SEQ ID NOS: 4
; SOFTWARE: FASTSEQ for Windows Version 4.0
; SEQ ID NO 1
; LENGTH: 2674
; TYPE: DNA
; ORGANISM: Human
US-09-817-180-1

Query Match 13.8%; Score 29; DB 4; Length 2674;
Best Local Similarity 54.1%; Pred. No. 2.2;
Matches 59; Conservative 0; Mismatches 50; Indels 0; Gaps 0;

QY 55 CTTAGGGAAGTCTGGGAAATCTCCCTGTTTGGGGGGGCGGCGGATGAC 114
DB 1083 CTCGGGATGAAGAGGACACCCCGGAGCGGGGTGAGCTGTGGCAGAGG 1142
QY 115 CAGGGCCGAGGAACTCTGAACTCCGTAGATTGCTGTAGCCGCC 163
DB 1143 CAGTCTCTCAGAGACACTGCAGGGGCTGCAAGTAGCCTGTGACGCC 1191

RESULT 2
PCT-US96-01735-5
; Sequence 5, Application PC/TUS9601735
; GENERAL INFORMATION:
; APPLICANT: Marks, Andrew R.
; TITLE OF INVENTION: HUMAN T CELL INOSITOL 1,4,5,-TRISPHOSPHATE
; TITLE OF INVENTION: RECEPTOR
; NUMBER OF SEQUENCES: 8
; CORRESPONDENCE ADDRESS:
ADDRESSEE: Brumbaugh, Graves, Donohue & Raymond
STREET: 30 Rockefeller Plaza
CITY: New York
COUNTRY: USA
STATE: NY
ZIP: 10112-0228
COMPUTER READABLE FORM:
MEDIUM TYPE: Diskette

COMPUTER: IBM Compatible
OPERATING SYSTEM: DOS
SOFTWARE: FastSeq Version 1.5
CURRENT APPLICATION DATA:
APPLICATION NUMBER: PCT/US96/01735
FILING DATE:
CLASSIFICATION:
PRIOR APPLICATION DATA:
APPLICATION NUMBER: 08/386,039
FILING DATE: 09-FEB-1995
ATTORNEY/AGENT INFORMATION:
NAME: KOLE, Lisa B.
REGISTRATION NUMBER: 35,225
REFERENCE/DOCKET NUMBER: A30042 - 165/30555
TELECOMMUNICATION INFORMATION:
TELEPHONE: 212-408-2628
TELEFAX: 212-765-2519
TELEX:
INFORMATION FOR SEQ ID NO: 5:
SEQUENCE CHARACTERISTICS:
LENGTH: 8791 base pairs
TYPE: nucleic acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: cDNA
HYPOTHETICAL: NO
ANTI-SENSE: NO
FRAGMENT TYPE:
ORIGINAL SOURCE:
FEATURE:
PCT-US96-01735-5

Query Match 13.5%; Score 28.4; DB 5; Length 8791;
Best Local Similarity 50.0%; Pred. No. 5.6;
Matches 71; Conservative 0; Mismatches 71; Indels 0; Gaps 0;

QY 30 AGGAGTCCGAGAACACAGCTACTTAGGAGGCTGGGAAAATCTCCCTGCTTT 89
DB 5431 AGGAGTCCGAGAACACAGCTACTTAGGAGGCTGGGAAAATCTCCCTGCTTT 89
QY 90 GGGGGGCGAGGGGGGGGATGAGCCGAGAGGCGGAGAACTCTGAAGACTCCGCTAGT 149
DB 5491 GGGTGAATGATGCTGCGCGAGGTTGCTGACCTTGACAGAGAGGGGCTTCCAACT 5550
QY 150 TGGCTAGACCGCTCTCAGACAC 171
DB 5551 AGTTATGACCTCATCATGAAC 5572

RESULT 3
US-08-949-155-49/c
Sequence 49, Application US/08949155
Patent No. 6271436
GENERAL INFORMATION:
APPLICANT: Pledrehita, Jorge A
APPLICANT: Bazer, Fuller W
TITLE OF INVENTION: Compositions and Methods for the
TITLE OF INVENTION: Generation of Transgenic Animal Species
NUMBER OF SEQUENCES: 51
CORRESPONDENCE ADDRESSES:
ADDRESSEE: ARNOLD, WHITE AND DURKEE
STREET: P.O. Box 4433
CITY: Houston
STATE: TX
COUNTRY: US
ZIP: 77210-4433
COMPUTER READABLE FORM:
MEDIUM TYPE: floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: PatentIn Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/949,155

FILING DATE: Concurrently Herewith
CLASSIFICATION: 800
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 60/027,338
FILING DATE: 11-OCT-1996
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 60/046,094
FILING DATE: 09-MAY-1997
ATTORNEY/AGENT INFORMATION:
NAME: Hibler, David W.
REGISTRATION NUMBER: 41,071
REFERENCE/DOCKET NUMBER: TAMK:177
TELECOMMUNICATION INFORMATION:
TELEPHONE: (512) 418-3000
TELEFAX: (713) 789-2679
INFORMATION FOR SEQ ID NO: 49:
SEQUENCE CHARACTERISTICS:
LENGTH: 4791 base pairs
TYPE: nucleic acid
STRANDEDNESS: single
TOPOLOGY: linear
US-08-949-155-49

Query Match 13.3%; Score 28; DB 4; Length 4791;
Best Local Similarity 55.0%; Pred. No. 6.1;
Matches 55; Conservative 0; Mismatches 45; Indels 0; Gaps 0;

QY 73 AAATCTCCCTGCTTTGGGGGCGAGGGCGGGGATGAGCCAGGCGGAGAACT 132
DB 193 AGATCTGCTGCTTTGGCGGAGGTTGACAAATGAGAGGGCGGAGAACT 134
QY 133 CTGAGACTCCGTAGATGCTGCTAGACCGCTCAGACT 172
DB 133 GTGAGACTAGGAGAACTGCTTGACATCCCTCCATT 94

RESULT 4
US-09-318-448-20/c
Sequence 20, Application US/09318448
Patent No. 6210950
GENERAL INFORMATION:
APPLICANT: Johnson, William G.
APPLICANT: Steenroos, Edward S.
TITLE OF INVENTION: METHODS FOR DIAGNOSING, PREVENTING, AND TREATING
FILE REFERENCE: 601-1-057
CURRENT APPLICATION NUMBER: US/09/318,448
NUMBER OF SEQ ID NOS: 46
SOFTWARE: PatentIn Ver. 2.0
SEQ ID NO 20
TYPE: DNA
LENGTH: 3228
ORGANISM: Homo sapiens
US-09-318-448-20

Query Match 13.2%; Score 27.8; DB 4; Length 3228;
Best Local Similarity 54.4%; Pred. No. 6.1;
Matches 56; Conservative 0; Mismatches 47; Indels 0; Gaps 0;

QY 89 TGGGGGCGAGGGGGGGGATGAGCCGAGAGGAGAACTCTGAAGACTCCGTAGA 148
DB 1689 TGGCTGTGAAGGGGGGTCAGAGAGGCCCAAGAGCCCTGAGGCTGTGAGA 1630
QY 149 TTGCTAGACCGCTCTCAGACACTCTGGGCGAGCGTGGAGAG 191
DB 1629 TGTGCTGACTCCTGCGACAGGAGGAGGGGAGG 1587

RESULT 5
US-08-785-310A-4/c
Sequence 4, Application US/08785310A
Patent No. 5840532

Db 1114 CTTTTCATTTCAGGCTTCCAGAGAACTTCTCAGAAAGAGTGCCTGATGAGGACAGG 1173
OY 65 GTCTGGAAATAATCTCCCTGCTTTGGGAGG 96
Db 1174 CTTGAGACACAGACACACACTTGAAGTGG 1205

RESULT 12
US-08-257-963B-10

; Sequence 10, Application US/08257963B
; Patent No. 5840866
; GENERAL INFORMATION:
; APPLICANT: Chader, Gerald J.; Becerra, S.
; APPLICANT: Patricia; Schwartz, Joan P.;
; APPLICANT: Tanikawa, Takayuki
; TITLE OF INVENTION: PIGMENT EPITHELIUM
; TITLE OF INVENTION: DERIVED FACTOR: CHARACTERIZATION OF ITS NOVEL
; TITLE OF INVENTION: BIOLOGICAL ACTIVITY AND SEQUENCES ENCODING
; TITLE OF INVENTION: AND EXPRESSING THE PROTEIN
; NUMBER OF SEQUENCES: 42
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Morgan & Flinnegan
; STREET: 345 Park Avenue
; CITY: New York
; STATE: New York
; COUNTRY: USA
; ZIP: 10154
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy Disk
; COMPUTER: IBM PC Compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: WORDPERFECT 5.1
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/257, 963B
; FILING DATE:
; CLASSIFICATION: 514
; PRIORITY APPLICATION DATA:
; APPLICATION NUMBER: 07/952,796
; FILING DATE: 24-SEPT-1992
; ATTORNEY/AGENT INFORMATION:
; NAME: DOROTHY R. AOTH
; REGISTRATION NUMBER: 36434
; REFERENCE/DOCKET NUMBER: 20264126US1
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (212) 758-4800
; TELEFAX: (212) 751-6849
; INFORMATION FOR SEQ ID NO: 10:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 7210 Base Pairs
; TYPE: Nucleic Acid
; STRANDEDNESS: Double
; TOPOLOGY: Unknown
; MOLECULE TYPE: Genomic DNA
; ORIGINAL SOURCE:
; ORGANISM: Human
; IMMEDIATE SOURCE:
; LIBRARY: DASH II
; FEATURE:
; NAME/KEY: JT106
; LOCATION:
; IDENTIFICATION METHOD:
; OTHER INFORMATION: 7.2 kb No. 5840866 1 fragments
; OTHER INFORMATION: Derived from human placental genomic DNA
US-08-257-963B-10

Query Match 12.5%; Score 26.2; DB 2; Length 7210;

Best Local Similarity 53.4%; Pred. No. 28;

Matches 55; Conservative 0; Mismatches 48; Indels 0; Gaps 0;

OY 22 GCCTGTCAAGAACTCCAGAAACACAGCTGACTTGAAGGAAAGTCTGGGAAAAATCTCC 81
Db 6869 GGCATCTCAGAGGACAGAAAAAGAGGGGTGCAAGAGAGAAATGCGAGACAGACAGCC 6928

OY 82 CTGCTTTTGGGGGCGAGGGCGGGGATGACCGAGGCCGAG 124
Db 6929 CTCGAATTTGGCGAAAAAGGTGATGATGAGAGAGGACAGAG 6971

RESULT 13
US-08-367-841A-10

; Sequence 10, Application US/08367841A
; Patent No. 6319687
; GENERAL INFORMATION:
; APPLICANT: Chader, Gerald J.; Rodriguez,
; APPLICANT: Ignacio R.; Mazuruk, Krzysztof;
; APPLICANT: Tombran-Tink, Joyce
; TITLE OF INVENTION: PIGMENT EPITHELIUM
; TITLE OF INVENTION: DERIVED FACTOR: CHARACTERIZATION GENOMIC
; TITLE OF INVENTION: ORGANIZATION AND SEQUENCE OF THE PEDF GENE
; NUMBER OF SEQUENCES: 43
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Morgan & Flinnegan
; STREET: 345 Park Avenue
; CITY: New York
; STATE: New York
; COUNTRY: USA
; ZIP: 10154
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy Disk
; COMPUTER: IBM PC Compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: WORDPERFECT 5.1
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/367, 841A
; FILING DATE: 30-DEC-1994
; CLASSIFICATION: 435
; PRIORITY APPLICATION DATA:
; APPLICATION NUMBER: 08/257, 963
; FILING DATE: 07-JUN-1994
; PRIORITY APPLICATION DATA:
; APPLICATION NUMBER: 07/952,796
; FILING DATE: 24-SEP-1992
; ATTORNEY/AGENT INFORMATION:
; NAME: DOROTHY R. AOTH
; REGISTRATION NUMBER: 36434
; REFERENCE/DOCKET NUMBER: 20264126US2
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (212) 758-4800
; TELEFAX: (212) 751-6849
; INFORMATION FOR SEQ ID NO: 10:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 7210 Base Pairs
; TYPE: Nucleic Acid
; STRANDEDNESS: Double
; TOPOLOGY: Unknown
; MOLECULE TYPE: Genomic DNA
; ORIGINAL SOURCE:
; ORGANISM: Human
; IMMEDIATE SOURCE:
; LIBRARY: DASH II
; FEATURE:
; NAME/KEY: JT6A
; LOCATION:
; IDENTIFICATION METHOD:
; OTHER INFORMATION: 7.0 kb No. 6319687 1-No. 6319687
; OTHER INFORMATION: fragment; Derived from human placental
; OTHER INFORMATION: genomic DNA; also referred to as JT106
US-08-367-841A-10

Query Match 12.5%; Score 26.2; DB 4; Length 7210;

Best Local Similarity 53.4%; Pred. No. 28;

Matches 55; Conservative 0; Mismatches 48; Indels 0; Gaps 0;

OY 22 GCCTGTCAAGAACTCCAGAAACACAGCTGACTTGAAGGAAAGTCTGGGAAAAATCTCC 81
Db 6869 GGCATCTCAGAGGACAGAAAAAGAGGGGTGCAAGAGAGAAATGCGAGACAGACAGCC 6928

Thu Nov 7 08:57:02 2002

us-09-833-740-7_copy_1_210.rni

Page 7

Qy 82 CTGCTTTTGGGGGGCAGGGCCGGGATGAGCCAGGGCCGAG 124
 ||||| - - - - - ||||| - ||||| ||||| ||||| |||||
 Db 6921 CTGCATTTTGGCAAAGGGTGAATGATGAGAGAGGCGCAGAG 6963

```
Search completed: November 6, 2002, 14:20:31
Job time : 42.9733 secs
```


GenCore version 5.1.3
Copyright (c) 1993 - 2002 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: November 6, 2002, 14:17:21 ; Search time 260.086 Seconds
(without alignments)
10897.801 Million cell updates/sec

Title: US-09-833-740-7_COPY_1_210
Perfect score: 210
Sequence: 1 ccgcctgtctctctcctc.....ggattgtgcaaacattcc 210

Scoring table: IDENTITY_NUC
Gapop 10.0 , Gapext 1.0

Searched: 13736207 seqs, 6748477542 residues

Total number of hits satisfying chosen parameters: 27472414

Minimum DB seq length: 0

Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%

Maximum Match 100%
Listing first 45 summaries

Database :

EST:*
1: em_estbta:*
2: em_estbta:*
3: em_estin:*
4: em_estinu:*
5: em_estov:*
6: em_estopl:*
7: em_estro:*
8: em_hic:*
9: gb_estl:*
10: gb_estl2:*
11: gb_hic:*
12: gb_gss:*
13: em_gss_hum:*
14: em_gss_inv:*
15: em_gss_pln:*
16: em_gss_vrti:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	52.6	25.0	689	12	AZ971837 2M0245B02
2	38.6	18.4	782	9	AU119486 AU119486
3	34.8	16.6	939	10	BG855275 1024042C0
4	34.2	16.3	771	10	BG479234 602526449
5	34	16.2	623	9	AA690752 vus7c08.r
6	33.8	16.1	895	10	BG755847 602716308
7	33.6	16.0	695	10	BI256085 602978917
8	33.6	16.0	836	12	AG076601 Pan trogl
9	33.2	15.8	258	10	BI831409 603074508
10	33.2	15.8	303	10	BF172864 MYE0121 M
11	33.2	15.8	545	12	CNS0172C Tetraodon
12	33.2	15.8	597	10	BE336705 ba99d01.y
13	33.2	15.8	662	10	BI828928 603075090
14	33.2	15.8	882	10	BG035198 60324706
15	33	15.7	386	10	BF759824 PMO-CT064
16	32.8	15.6	921	12	AQ898422 RS_315_B
17	32.8	15.6	1101	12	CNS017R9 Drosophila

C	18	32.6	15.5	430	9	BB776545	BB776545
	19	32.6	15.5	490	9	AI561790	AI561790 vv6b12.x
	20	32.6	15.5	1019	12	CNS05MS	AL43657 Tetraodon
	21	32.4	15.4	288	9	AA788598	AA788598 ag58q10.s
	22	32.4	15.4	300	9	AW517011	AW517011 xp90c11.x
	23	32.4	15.4	334	10	BF001888	BF001888 7994h12.x
	24	32.4	15.4	356	10	BF195077	BF195077 7094f04.x
	25	32.4	15.4	377	9	AA776500	AA776500 ah11f10.s
	26	32.4	15.4	380	9	AA439335	AA439335 xc17g05.x
	27	32.4	15.4	407	9	AI874207	AI874207 km50c02.x
	28	32.4	15.4	411	9	AA473156	AA473156 xy14h02.x
	29	32.4	15.4	412	10	BM127741	BM127741 1f08c08.x
	30	32.4	15.4	413	9	AW169135	AW169135 xj19a09.x
C	31	32.4	15.4	428	10	BM127990	BM127990 1f08c08.y
	32	32.4	15.4	447	9	AI804686	AI804686 lt94e04.x
	33	32.4	15.4	461	9	AI184928	AI184928 ok28c02.x
	34	32.4	15.4	464	9	AI201132	AI201132 qf64f09.x
	35	32.4	15.4	502	9	AA779225	AA779225 zj39e07.s
	36	32.4	15.4	567	9	AA543126	AA543126 vk36b12.r
	37	32.4	15.4	574	10	BI438774	BI438774 1c26e06.x
C	38	32.4	15.4	674	12	AG002186	AG002186 Homo sapi
	39	32.4	15.4	687	12	AG002187	AG002187 Homo sapi
C	40	32.4	15.4	707	12	AG002182	AG002182 Homo sapi
	41	32.4	15.4	750	12	AG002170	AG002170 Homo sapi
	42	32.4	15.4	755	12	AG002169	AG002169 Homo sapi
	43	32.4	15.4	777	10	BI463395	BI463395 60320431
C	44	32.2	15.3	305	9	AA153765	AA153765 m04b03.r
	45	32.2	15.3	519	9	AU175395	AU175395 AU175395

ALIGNMENTS

RESULT 1
LOCUS AZ971837 689 bp DNA linear GSS 27-Apr-2001
DEFINITION 2M0245B02R Mouse 10kb plasmid UUGC2M library Mus musculus genomic
clone UUGC2M0245B02 R, DNA sequence.
ACCESSION AZ971837
VERSION AZ971837.1 GI:1843064
KEYWORDS GSS.
SOURCE house mouse.
ORGANISM Mus musculus
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
REFERENCE 1 (bases 1 to 689)
Dunn,D., Aoyagi,A., Barber,M., Beacorn,T., Duval,B., Hamil,C.,
Islam,H., Longacre,S., Mahmoud,M., Meenen,E., Pedersen,T., Reilly,
M., Rose,M., Rose,R., Stokes,R., Tingey,A., von Niederhausern,A.
and Wright,D., Weiss,R.
Mouse whole genome scaffolding with paired end reads from 10kb
plasmid inserts
JOURNAL Unpublished (2000)
COMMENT Contact: Robert B. Weiss
University of Utah
Genome Center
Rm. 308, Biomedical Polymers Research Bldg., 20 S. 2030 E., SLC, UT
84112, USA
Tel: 801 585 5606
Fax: 801 585 7177
Email: ddunn@genetics.utah.edu
Insert Length: 1000 Std Error: 0.00
Plate: 0245 row: B column: 02
Seq primer: CACACAGGAACACGATATGACC
Class: Plasmid ends
High quality sequence stop: 689.
Location/Qualifiers
1..689
/organism="Mus musculus"
/strain="C57BL/6J"
/db_xref="taxon:10090"
/clone="UUGC2M0245B02"
/clone_lib="Mouse 10kb plasmid UUGC2M library"

FEATURES

source

/sex="Female"
/lab host="E. coli strain XL10-Gold, T1-resistant, F-"
/note="Vector: pMD22ny: Purified genomic DNA from M.
musculus C57BL/6J (female) was obtained from the Jackson
laboratory Mouse DNA Resource
(<http://www.jax.org/resources/documents/dnares/>). The DNA
was hydrodynamically sheared by repeated passage through a
0.005 inch orifice at constant velocity. The sheared DNA
was blunt end-repaired with T4 DNA polymerase and T4
polynucleotide kinase. Adaptor oligonucleotides were
ligated to the blunt ends in high molar excess. The
adaptor DNA was purified and size-selected for a 9.5 to
10.5 kb range using preparative agarose gel
electrophoresis. Vector DNA was prepared from a derivative
of pMD22 (g11473211419b1fl21072.1), a copy-number
inducible derivative of plasmid R1. The vector was ligated
with adaptors complementary to the insert adaptors and
purified. The sheared, adaptor mouse DNA was annealed to
chemically-competent E. coli XL10-Gold (Stratagene) cells
and selected for ampicillin resistance."

BASE COUNT	173 a	189 c	147 g	180 t
ORIGIN				

Query Match	25.0%;	Score 52.6;	DB 12;	Length 689;
Best Local Similarity	68.5%;	Pred. No. 0.00013;		
Matches 87; Conservative	0;	Mismatches 39;	Indels 1;	Gaps 1;

Oy 3 GCCCTGTTCTTCTTCCTCAGCCGTGCAGGAGTCCAGAAAGCAGCTGACTT-AGGG 61
 1 111 1111 1 11 1111111111 11 1 111 111
 Db 137 GCTTCGCTTTTCTTCGCGGCTGCTGAGGAAGTCCAGGACAGCTAGACGCTTGGGG 78

Oy 62 AAGCTCTGGGAAAAATCTCCCTGCTTTTGGGGGGGCGAGGCGGGGAGTGAAGCCAGGCCC 12
 |||||
 Db 77 TAGGCTCTGGGAAAAATCTCCCAAGATTTTAAAGAGGGGCGAGCGGGGCGATGACAAACTGGA 18

QY	122	GAGAGG	128
Db	17	GATTCG	11

RESULT	2
AU119486/c	
LOCUS	AU119486 HEMBA1 Homo sapiens cDNA clone HEMBAL005931 5', mRNA
DEFINITION	772 bp linear EST 19-OCT-2000
ACCESSION	AU119486
VERSION	AU119486
KEYWORDS	sequence.
SOURCE	AU119486.1 GI:10934721
	EST.
	human.

REFERENCE
Ota, T., Nishikawa, T., Suzuki, Y., Ishii, S., Saito, K., Kawai, Y.,
1 (bases 1 to 772)
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominiidae; Homo.

TITLE	HRI human cDNA project
JOURNAL	Unpublished (2000)
COMMENT	Contact: Takao Isogai

Helix Research Institute
1532-3 Yama, Kisarazu, Chiba 292-0812, Japan
Tel.: 81-438-52-3951
Fax: 81-438-52-3952
Email: genomics@hri.co.jp
HRI human cDNA project: 5'- & 3'-end one pass sequencing; Helix
Research Institute; cDNA library construction; Department of
Virology, Institute of Medical Science, University of Tokyo, and
Helix Research Institute.

FEATURES	Location/Qualifiers
source	1. .772
	/organism="Homo sapiens"

BASE COUNT	ORIGIN
147 a	258 c 194 g 170 t 3 others

Query Match	18.4%;	Score 38.6;	DB 9;	Length 772;
Best Local Similarity	52.1%;	Pred. No. 1.3;		
Matches	86;	Conservative	0;	Mismatches 79;
				Indels 0;
				Gaps 0;

QY 37 CCAGGAACGACAGCTGACTTAGGAGAGCTCTGGAAATAATCCTCCGTCTTTGGGGGG 96
 ||||| ||||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 258 CCCAAGAAGCACAGGACCACCTTGGGAAAAGACAGGGGACAGTCTTCACAGCTTGCGACCCCT 139

Qy 97 CAGGGGCGGGGGATGACCCAGGGCCGAGAAGGAACCTGTGAAGACTCCGTAGATTGCTCTTA 156
| | | | |
Db 198 TAAATCTGAGAGAGAGAAGCTCTTTGCCCTGGGGGAACATTACAGAGCCTGGAGCATAGSSTTA 139

QY 157 GACCGCCTCAGACACTCTCGGGCGAGCGTGAGAGGATTTGTCA 201
||| | | | | | | | | | | | | | |
Db 138 CACCTCATAGGAGAGACAGGCCCTGATCCTGGTGAGGGCTTGGCCA 94

RESULT 3	LOCUS	DEFINITION
CG855275/c	CG855275	939 bp mRNA linear EST 29-MAY-2001
	1024042C07.y1 C.	reinhardtii CC-1690, normalised, Lambda Zap II
	Chlamydomonas reinhardtii cDNA,	mRNA sequence.

ACCESSION	BGR55275	GI:14236459
VERSION	BGR55275.1	
KEYWORDS	EST.	
SOURCE	Chlamydomonas reinhardtii.	
ORGANISM	Chlamydomonas reinhardtii	

REFERENCE
AUTHORS
Grossman, A., Davies, J., Federspiel, N., Harris, E., Lefebvre, P.,
1 (bases 1 to 939)
Chlamydomonadaceae; Chlamydomonas.
Chlamydomonadaceae; Chlamydomonas.

TITLE	Analyses of the <i>Chlamydomonas reinhardtii</i> Genome: A Model, Unicellular System for Analyzing Gene Function and Regulation in Vascular Plants: project phase 2
JOURNAL	Unpublished (2000)
COMMENT	Contact: Charles Hauser

Duke University
Durham, NC 27708-1000
Tel: 919 613 8159
Fax: 919 613 8177
Email: chauser@duke.edu

FEATURES	Location/Qualifiers
source	1. . 939

```

/organism="Chlamydomonas reinhardtii"
/strain="CC-1690 wild type mt+ 21sr"
/db_xref="taxon:3055"
/clone_id="C. reinhardtii CC-1690, normalized, Lambda Zap
II"

```

```

/note=Vector: plusescript II SK-; Site_1: EcoRI; Site_2:
XhoI; This library, constructed by John Davies and Jeffrey
McMormott, combines cDNAs from CC-1690 cells grown to
mid-log phase in RAP (acetate-containing) medium in the
light. TAP medium in the dark. HS (minimim) medium in

```

light, 10% medium in the dark, 45 (turning) medium, 10% ambient levels of CO₂ and 45 medium bubbled with 5% CO₂. PolyA mRNA was purified from each sample, pooled and cDNA synthesized. The cDNA was directionally cloned into lambda Zap II (Stratagene) in the EcoRI (5') and XhoI (3') sites. Characteristic cDNA bands were purified from the lambda

BASE COUNT 199 a 336 c 140 g 264 t

Db 228 GGGTAAGCCAAAGACGAGGGGTAGGATTCAGATTCAGACTGGTGTAC 175

RESULT 6

Bg755847

LOCUS

Bg755847

DEFINITION

602776308F1 NIH_MGC_48 Homo sapiens cDNA clone IMAGE:485665 5', mRNA sequence.

ACCESSION

Bg755847

VERSION

Bg755847.1 GI:14066500

KEYWORDS

EST.

SOURCE

human.

ORGANISM

Homo sapiens

REFERENCE 1 (bases 1 to 895)
 NIH-MGC http://mhc.nci.nih.gov/
 National Institutes of Health, Mammalian Gene Collection (MGC)
 Unpublished (1999)
 Contact: Robert Strausberg, Ph.D.
 Email: cga@bbs-remail.nih.gov
 Tissue Procurement: ATCC
 cDNA Library Preparation: Life Technologies, Inc.
 DNA Sequencing by: Incyte Genomics, Inc.
 Clone distribution: MGC clone distribution information can be found through the I.M.A.G.E. Consortium/LNL at:
 http://image.llnl.gov
 Plate: L1302 row: b column: 06
 High quality sequence stop: 5
 Location/Qualifiers

FEATURES

source

1. 895

/organism="Homo sapiens"

/db_xref="taxon:9606"

/clone="IMAGE:485665"

/tissue="NIH_MGC_48"

/lab_host="DH10B (phage-resistant)"

/note="Organ: B-cells from tonsils (cell line)"

/site="EcoRI; CDNA made by oligo-dT priming. Directionally cloned into EcoRI/XhoI sites using the following 5' adaptor: GGCACG(6). Size-selected >500bp for average insert size 1.8kb. Library constructed by Ling Hong in the laboratory of Gerald M. Rubin (University of California, Berkeley) using ZAP-cDNA synthesis kit (Stratagene) and Superscript II RT (Life Technologies). Note: This is a NIH-MGC Library."

BASE COUNT

177 a 284 c 292 g 141 t

ORIGIN

Query Match

Best Local Similarity 56.4%; Pred. No. 31;

Matches 62; Conservative 0; Mismatches 48; Indels 0; Gaps 0;

16.1%; Score 33.8; DB 10; Length 895;
 56.4%; Pred. No. 31;
 Matches 62; Conservative 0; Mismatches 48; Indels 0; Gaps 0;

RESULT 7

Bg755847

LOCUS

Bg755847

DEFINITION

602978917F1 NCI_CGAP_L19 Mus musculus cDNA clone IMAGE:5123621 5', mRNA sequence.

ACCESSION

Bg755847

VERSION

Bg755847.1 GI:14810144

KEYWORDS

EST.

SOURCE

house mouse.

ORGANISM

Mus musculus

REFERENCE 1 (bases 1 to 695)
 NIH-MGC http://mhc.nci.nih.gov/
 National Institutes of Health, Mammalian Gene Collection (MGC)
 Unpublished (1999)
 Contact: Robert Strausberg, Ph.D.
 Email: cga@bbs-remail.nih.gov
 Tissue Procurement: ATCC
 cDNA Library Preparation: Life Technologies, Inc.
 DNA Sequencing by: Incyte Genomics, Inc.
 Clone distribution: MGC clone distribution information can be found through the I.M.A.G.E. Consortium/LNL at:
 http://image.llnl.gov
 Plate: L1302 row: b column: 06
 High quality sequence stop: 5
 Location/Qualifiers

FEATURES

source

1. 695

/organism="Mus musculus"

/db_xref="taxon:10090"

/clone="IMAGE:5123621"

/tissue="NCI_CGAP_L19"

/lab_host="DH10B (T1 phage-resistant)"

/note="Organ: Liver; Vector: pCMV-SPORT6; Site 1: NotI; Site 2: SalI; Cloned unidirectionally. Primer: Oligo dT. Average insert size 1.9 kb. Constructed by Life Technologies. Note: this is a NCI_CGAP Library."

BASE COUNT

148 a 223 c 187 g 137 t

ORIGIN

Query Match

Best Local Similarity 53.9%; Pred. No. 33;

Matches 69; Conservative 0; Mismatches 59; Indels 0; Gaps 0;

16.0%; Score 33.6; DB 10; Length 695;
 53.9%; Pred. No. 33;
 Matches 69; Conservative 0; Mismatches 59; Indels 0; Gaps 0;

RESULT 8

AG076601

LOCUS

AG076601

DEFINITION

Pan troglodytes DNA, clone: PTB-070N17.F, genomic survey sequence.

ACCESSION

AG076601

VERSION

AG076601.1 GI:16628403

KEYWORDS

GSS (genome survey sequence).

SOURCE

Pan troglodytes male lymphoblast DNA, clone: PTB-070N17.F.

ORGANISM

Pan troglodytes

REFERENCE 1 (sites)
 Fujiyama, A., Hattori, M., Toyoda, A., Taylor, T.D., Yada, T., Totoki, Y., Watanabe, H. and Sakaki, Y.
 BAC end sequences of library PTB
 Unpublished
 2 (bases 1 to 836)
 Fujiyama, A., Hattori, M., Toyoda, A., Taylor, T.D., Yada, T., Totoki, Y., Watanabe, H. and Sakaki, Y.
 Direct Submission
 Submitted (02-AUG-2001) Asao Fujiyama, The Institute of Physical

RESULT 8

AG076601

LOCUS

AG076601

DEFINITION

Pan troglodytes DNA, clone: PTB-070N17.F, genomic survey sequence.

ACCESSION

AG076601

VERSION

AG076601.1 GI:16628403

KEYWORDS

GSS (genome survey sequence).

SOURCE

Pan troglodytes male lymphoblast DNA, clone: PTB-070N17.F.

ORGANISM

Pan troglodytes

REFERENCE 1 (sites)
 Fujiyama, A., Hattori, M., Toyoda, A., Taylor, T.D., Yada, T., Totoki, Y., Watanabe, H. and Sakaki, Y.
 BAC end sequences of library PTB
 Unpublished
 2 (bases 1 to 836)
 Fujiyama, A., Hattori, M., Toyoda, A., Taylor, T.D., Yada, T., Totoki, Y., Watanabe, H. and Sakaki, Y.
 Direct Submission
 Submitted (02-AUG-2001) Asao Fujiyama, The Institute of Physical

RESULT 8

AG076601

LOCUS

AG076601

DEFINITION

Pan troglodytes DNA, clone: PTB-070N17.F, genomic survey sequence.

ACCESSION

AG076601

VERSION

AG076601.1 GI:16628403

KEYWORDS

GSS (genome survey sequence).

SOURCE

Pan troglodytes male lymphoblast DNA, clone: PTB-070N17.F.

ORGANISM

Pan troglodytes

REFERENCE 1 (sites)
 Fujiyama, A., Hattori, M., Toyoda, A., Taylor, T.D., Yada, T., Totoki, Y., Watanabe, H. and Sakaki, Y.
 BAC end sequences of library PTB
 Unpublished
 2 (bases 1 to 836)
 Fujiyama, A., Hattori, M., Toyoda, A., Taylor, T.D., Yada, T., Totoki, Y., Watanabe, H. and Sakaki, Y.
 Direct Submission
 Submitted (02-AUG-2001) Asao Fujiyama, The Institute of Physical

RESULT 8

AG076601

LOCUS

AG076601

DEFINITION

Pan troglodytes DNA, clone: PTB-070N17.F, genomic survey sequence.

ACCESSION

AG076601

VERSION

AG076601.1 GI:16628403

KEYWORDS

GSS (genome survey sequence).

SOURCE

Pan troglodytes male lymphoblast DNA, clone: PTB-070N17.F.

ORGANISM

Pan troglodytes

REFERENCE 1 (sites)
 Fujiyama, A., Hattori, M., Toyoda, A., Taylor, T.D., Yada, T., Totoki, Y., Watanabe, H. and Sakaki, Y.
 BAC end sequences of library PTB
 Unpublished
 2 (bases 1 to 836)
 Fujiyama, A., Hattori, M., Toyoda, A., Taylor, T.D., Yada, T., Totoki, Y., Watanabe, H. and Sakaki, Y.
 Direct Submission
 Submitted (02-AUG-2001) Asao Fujiyama, The Institute of Physical

RESULT 8

AG076601

LOCUS

AG076601

DEFINITION

Pan troglodytes DNA, clone: PTB-070N17.F, genomic survey sequence.

ACCESSION

AG076601

VERSION

AG076601.1 GI:16628403

KEYWORDS

GSS (genome survey sequence).

SOURCE

Pan troglodytes male lymphoblast DNA, clone: PTB-070N17.F.

ORGANISM

Pan troglodytes

REFERENCE 1 (sites)
 Fujiyama, A., Hattori, M., Toyoda, A., Taylor, T.D., Yada, T., Totoki, Y., Watanabe, H. and Sakaki, Y.
 BAC end sequences of library PTB
 Unpublished
 2 (bases 1 to 836)
 Fujiyama, A., Hattori, M., Toyoda, A., Taylor, T.D., Yada, T., Totoki, Y., Watanabe, H. and Sakaki, Y.
 Direct Submission
 Submitted (02-AUG-2001) Asao Fujiyama, The Institute of Physical

RESULT 8

AG076601

LOCUS

AG076601

DEFINITION

Pan troglodytes DNA, clone: PTB-070N17.F, genomic survey sequence.

ACCESSION

AG076601

VERSION

AG076601.1 GI:16628403

KEYWORDS

GSS (genome survey sequence).

SOURCE

Pan troglodytes male lymphoblast DNA, clone: PTB-070N17.F.

ORGANISM

Pan troglodytes

COMMENT

and Chemical Research (RIKEN), Genomic Sciences Center (GSC);
1-7-22 Suehiro-chou, Tsurumi-ku, Yokohama, Kanagawa 230-0045, Japan
(E-mail: chimbesc@sc.riken.go.jp, URL: http://hgp.gsc.riken.go.jp/
Tel: 81-45-503-9111, Fax: 81-45-503-9170)
Clones are derived from the chimpanzee BAC library PTB This BAC end
was generated during the R&D process and may have higher chance of
clone tracking errors.

FEATURES

Sequencing: -21M13
LIBRARY
Vector : pKS145
R.Site 1 : SacI
R.Site 2 : SacI
Location/Qualifiers
1. .836

BASE COUNT 145 a 178 c 237 g 261 t 15 others
ORIGIN

Query Match 16.0%; Score 33.6; DB 12; Length 836;
Best Local Similarity 51.3%; Pred. No. 35;
Matches 78; Conservative 0; Mismatches 74; Indels 0; Gaps 0;

OY 48 CAGCTGACTTAGGAGGAGCTGGGAAATCTCCCTTTTGGGGGCGGAGGGGGG 107
DB 253 CGGCAAGTCTGTGGTGGAGCCCAATCTTCCCGACAGGGGGGGGTGTGC 312
OY 108 GATGAGCCAGGGCCGAGAGAACTCTGAGACTCCGTGATTCCTGACCGCTCAG 167
DB 313 GGTTCGGGGGGGGTACGATTAATGTTGCTTGTGTGGTTAAACCGCCAAAT 372
OY 168 ACATCTCGGGCGGCGAGCGAGAGATTTGTG 199
DB 373 TACTCTACTACCCAGTGTGTGGGCTGTGTG 404

RESULT 9
BI831409/c 258 bp mRNA linear EST 04-OCT-2001
LOCUS 603074508F1 NIH_MGC_119 Homo sapiens cDNA clone IMAGE:5166577 5',
DEFINITION mRNA sequence.
ACCESSION BI831409
VERSION BI831409.1 GI:15942959
KEYWORDS EST.
SOURCE human.
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
REFERENCE 1 (bases 1 to 258)
AUTHORS NIH-MGC http://mgc.nci.nih.gov/
TITLE National Institutes of Health, Mammalian Gene Collection (MGC)
JOURNAL Unpublished (1999)
COMMENT Contact: Robert Strausberg, Ph.D.
Email: c9apds-femail.nih.gov
Tissue Procurement: Life Technologies, Inc.
CDNA Library Preparation: Life Technologies, Inc.
CDNA Library Arrayed by: The I.M.A.G.E. Consortium (LNL)
DNA Sequencing by: Incyte Genomics, Inc.
Clone distribution: MGC clone distribution information can be
found through the I.M.A.G.E. Consortium/LNL at:
http://image.llnl.gov
Plate: L14M1413 row: P column: 02
High quality sequence stop: 258.
Location/Qualifiers
1. .258

FEATURES

source
/organism="Homo sapiens"
/db_xref="taxon:9606"
/clone="IMAGE:5166577"

BASE COUNT 45 a 97 c 69 g 47 t
ORIGIN

Query Match 15.8%; Score 33.2; DB 10; Length 258;
Best Local Similarity 67.1%; Pred. No. 31;
Matches 47; Conservative 0; Mismatches 23; Indels 0; Gaps 0;

OY 59 GGAAGTCTGGGAAATCTCCCTTTTGGGGGCGGAGGGGGGATGACCCAGG 118
DB 209 GGGAGGACGTGGGCAATGGCCCTCTGCTGGGTGAATAGTGGTGTAAAGCCAGG 150
OY 119 GCCGAGAGG 128
DB 149 GCCGGGAGAG 140

RESULT 10
BF172864 303 bp mRNA linear EST 23-MAR-2001
LOCUS BF172864
DEFINITION MYE0121 Myeloma (MYE) cDNA library Homo sapiens cDNA, mRNA
sequence.
ACCESSION BF172864
VERSION BF172864.1 GI:13439132
KEYWORDS EST.
SOURCE human.
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
REFERENCE 1 (bases 1 to 303)
AUTHORS Claudio J.O., Tang H., Khan E.M., Voralia M., Li Z., Cukerman E.,
Francisco-Pabalan O., Ilew C.C. and Stewart A.K.
TITLE The transcriptional phenotype of myeloma cells
JOURNAL Unpublished (2000)
COMMENT Contact: A. Keith Stewart, M.D.
Oncology Research
University Health Network
610 University Ave., 5-126, Toronto, Ontario, M5G 2M9, Canada
Tel: (416) 946-4639
Fax: (416) 946-6546
Email: k.stewart@utoronto.ca
PCR PRIMERS
FORWARD: 5'-GCCAAGTCGAATTACCTCCTCACTAAGG-3'
BACKWARD: 5'-CGAATGATGTATAGACCTCACTAAGGCG-3'
Seq primer: 5'-GAATTAACCTCCTCACTAAGG-3'.
Location/Qualifiers
1. .303

FEATURES

source

/organism="Homo sapiens"
/db_xref="taxon:9606"
/clone="lib="Myeloma (MYE) cDNA library"
/sex="male"
/tissue-type="Blood"
/cell_type="myeloma"
/dev-stage="multiple myeloma"
/note="Vector: Lambda Zap Express; site: 1: EcoRI; site: 2:
XhoI; Myeloma cells from multiple myeloma patients' bone
marrow were purified by magnetic cell sorting. mRNA were
purified and an oligo d(T)18 primer containing XhoI
restriction site was used to prime first strand synthesis
using M-MV reverse transcriptase. To protect the cDNAs
from XhoI digestion in subsequent cloning step, the
nucleotide analogue 5-methyl-dCTP was added to the

RESULT 13
 BI828928 662 bp mRNA linear EST 04-OCT-2001
 LOCUS 603075090F1 NIH_MGC_119 Homo sapiens cDNA clone IMAGE:5166594 5',
 DEFINITION mRNA sequence.
 ACCESSION BI828928
 VERSION BI828928.1 GI:15940478
 KEYWORDS EST.
 SOURCE human.
 ORGANISM Homo sapiens
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 Mammalia; Eutheria; Primates; Catarrhini; Hominiidae; Homo.
 REFERENCE 1 (bases 1 to 662)
 NIH-MGC <http://imgc.nci.nih.gov/>.
 TITLE National Institutes of Health, Mammalian Gene Collection (MGC)
 JOURNAL Unpublished (1999)
 COMMENT Contact: Robert Strausberg, Ph.D.
 Email: cgapbs-remail.nih.gov
 Tissue Procurement: Life Technologies, Inc.
 cDNA Library Preparation: Life Technologies, Inc.
 DNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
 DNA Sequencing by: Incyte Genomics, Inc.
 Clone distribution: MGC clone distribution information can be
 found through the I.M.A.G.E. Consortium/LLNL at:
<http://image.llnl.gov>
 Plate: LHAM1414 row: d column: 23
 High quality sequence stop: 662.
 Location/Qualifiers
 1. 662
 /organism="Homo sapiens"
 /db_xref="taxon:9606"
 /clone_image="IMAGE:5166594"
 /clone_lib="NIH_MGC_119"
 /tissue_type="medulla"
 /lab_host="DH10B"
 /note="Organ: brain; Vector: pCMV-SPORT6; Site_1: NotI;
 Site_2: EcoRV (destroyed); RNA source normal medulla from
 anonymous male age 27. Library is oligo-dT primed and
 directionally cloned (EcoRV site is destroyed upon
 cloning). Average insert size 1.3 kb, insert size range
 0.9-3 kb. Library is normalized and enriched for
 full-length clones and was constructed by C. Gruber
 (Invitrogen). Research Genetics tracking code 013. Note:
 this is a NIH_MGC Library."
 BASE COUNT 111 a 227 c 203 g 121 t
 ORIGIN
 Query Match 15.8%; Score 33.2; DB 10; Length 662;
 Best Local Similarity 67.1%; Pred. No. 42;
 Matches 47; Conservative 0; Mismatches 23; Indels 0; Gaps 0;
 QY 59 GCGAAGTCTGGGAAATCTCCCTCTTTGGGGGCGGCGGATGAGCCAGG 118
 ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
 Db 613 GGGAGGACCTGGGCAATGGCCCCCTCTTGGGTGGAATAGTGGTGAATAAGGCGAGG 554
 QY 119 GCCGGAAGG 128
 ||||| |||||
 Db 553 GCCGGGAGG 544
 RESULT 14
 BG035198 862 bp mRNA linear EST 24-JAN-2001
 LOCUS 603324706F1 NIH_MGC_90 Homo sapiens cDNA clone IMAGE:4413048 5',
 DEFINITION mRNA sequence.
 ACCESSION BG035198
 VERSION BG035198.1 GI:12429091
 KEYWORDS EST.
 SOURCE human.
 ORGANISM Homo sapiens
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;

Mammalia; Eutheria; Primates; Catarrhini; Hominiidae; Homo.
 REFERENCE 1 (bases 1 to 862)
 NIH-MGC <http://imgc.nci.nih.gov/>.
 TITLE National Institutes of Health, Mammalian Gene Collection (MGC)
 JOURNAL Unpublished (1999)
 COMMENT Contact: Robert Strausberg, Ph.D.
 Email: cgapbs-remail.nih.gov
 Tissue Procurement: ATCC
 cDNA Library Preparation: Life Technologies, Inc.
 DNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
 DNA Sequencing by: Incyte Genomics, Inc.
 Clone distribution: MGC clone distribution information can be
 found through the I.M.A.G.E. Consortium/LLNL at:
<http://image.llnl.gov>
 Plate: LHAM10138 row: k column: 01
 High quality sequence stop: 690.
 Location/Qualifiers
 1. 862
 /organism="Homo sapiens"
 /db_xref="taxon:9606"
 /clone_image="IMAGE:4413048"
 /clone_lib="NIH_MGC_90"
 /tissue_type="adenocarcinoma, cell line"
 /lab_host="DH10B (phage-resistant)"
 /note="Organ: liver; Vector: pCMV-SPORT6; Site_1: NotI;
 Site_2: SalI; Cloned unidirectionally; oligo-dT primed.
 Average insert size 1.7 kb. Library enriched for
 full-length clones and constructed by Life Technologies.
 Note: this is a NIH_MGC Library."
 BASE COUNT 197 a 308 c 215 g 162 t
 ORIGIN
 Query Match 15.8%; Score 33.2; DB 10; Length 862;
 Best Local Similarity 67.1%; Pred. No. 46;
 Matches 47; Conservative 0; Mismatches 23; Indels 0; Gaps 0;
 QY 59 GCGAAGTCTGGGAAATCTCCCTCTTTGGGGGCGGCGGATGAGCCAGG 118
 ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
 Db 387 GGGAGGACCTGGGCAATGGCCCCCTCTTGGGTGGAATAGTGGTGAATAAGGCGAGG 328
 QY 119 GCCGGAAGG 128
 ||||| |||||
 Db 327 GCCGGGAGG 318
 RESULT 15
 BF759824 386 bp mRNA linear EST 12-JAN-2001
 LOCUS PMO-CT0642-151200-001-g01 CT0642 Homo sapiens cDNA, mRNA sequence.
 DEFINITION BF759824
 ACCESSION BF759824
 VERSION BF759824.1 GI:12107724
 KEYWORDS EST.
 SOURCE human.
 ORGANISM Homo sapiens
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 Mammalia; Eutheria; Primates; Catarrhini; Hominiidae; Homo.
 REFERENCE 1 (bases 1 to 386)
 Dias Neto, E., Garcia Correa, R., Verjovski-Almeida, S., Briones, M.R.,
 Nagai, M.A., da Silva, W. Jr., Zago, M.A., Bordin, S., Costa, F.F.,
 Goldman, G.H., Carvalho, A.F., Matsukuma, A., Bala, G.S., Simpson, D.H.,
 Brunstein, A., deOliveira, P.S., Bucher, P., Jongeneel, C.V., O'Hare,
 M.J., Soares, F., Brentani, R.R., Reis, L.F., de Souza, S.J. and
 Simpson, A.J.
 Shotgun sequencing of the human transcriptome with ORF expressed
 sequence tags
 Proc. Natl. Acad. Sci. U.S.A. 97 (7), 3491-3496 (2000)
 Contact: Simpson A.J.G.
 Laboratory of Cancer Genetics
 Ludwig Institute for Cancer Research
 Rua Prof. Antonio Prudente 109, 4 andar, 01509-010, Sao Paulo-SP,
 Brazil
 Tel: +55-11-2704922

GenCore version 5.1.3
Copyright (c) 1993 - 2002 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: November 6, 2002, 14:05:40 : Search time 63.0267 Seconds
(without alignments)
6469.302 Million cell updates/sec

Title: US-09-833-740-1_COPY_1_1660
Perfect score: 1660
Sequence: 1 aggttaccgcatttgaca.....ggaagagacttgcaaca 1660

Scoring table: IDENTITY_NUC
Gapop 10.0, Gapext 1.0

Searched: 383533 seqs, 122816752 residues

Total number of hits satisfying chosen parameters: 767066

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : Issued_Patents_NA :
1: /cgn2_6/ptodata/1/ina/5A.COMB.seq.*
2: /cgn2_6/ptodata/1/ina/5B.COMB.seq.*
3: /cgn2_6/ptodata/1/ina/6A.COMB.seq.*
4: /cgn2_6/ptodata/1/ina/6B.COMB.seq.*
5: /cgn2_6/ptodata/1/ina/PCNUS.COMB.seq.*
6: /cgn2_6/ptodata/1/ina/backfile1.seq.*

Pred. No. is the number of results predicted by chance to have a
score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	58	3.5	3066	1 US-08-142-439A-1	Sequence 1, Appli
2	58	3.5	3066	2 US-08-869-477-1	Sequence 1, Appli
3	58	3.5	3381	4 US-09-009-119-1	Sequence 1, Appli
4	58	3.5	3381	4 US-09-371-507-1	Sequence 1, Appli
5	58	3.5	3383	5 PCT-US95-09098-1	Sequence 1, Appli
6	58	3.5	10409	3 US-08-772-440-3	Sequence 33, Appli
7	57.6	3.5	200	1 US-08-222-177A-8	Sequence 8, Appli
8	57.4	3.5	48974	4 US-08-920-422-17	Sequence 17, Appli
9	57.2	3.4	372	3 US-08-750-064-1	Sequence 1, Appli
10	57.2	3.4	372	3 US-08-545-196B-14	Sequence 14, Appli
11	57.2	3.4	80246	4 US-08-222-177A-14	Sequence 4, Appli
12	57	3.4	223	1 US-08-720-484A-1	Sequence 1, Appli
13	56.4	3.4	3854	3 US-08-953-823A-1	Sequence 14, Appli
14	56.4	3.4	7218	1 US-08-232-463-14	Sequence 14, Appli
15	56.4	3.4	818	4 US-08-702-525-4	Sequence 4, Appli
16	56.2	3.4	818	4 US-08-702-525-4	Sequence 4, Appli
17	56.2	3.4	818	5 PCT-US95-02576-4	Sequence 4, Appli
18	56.2	3.4	1570	4 US-08-205-657A-10	Sequence 10, Appli
19	56.2	3.4	1570	4 US-08-702-525-10	Sequence 10, Appli
20	56.2	3.4	1570	5 PCT-US95-02576-10	Sequence 10, Appli
21	56.2	3.4	1606	4 US-08-702-525-64	Sequence 64, Appli
22	56.2	3.4	1606	5 PCT-US95-02576-64	Sequence 64, Appli
23	56.2	3.4	1888	4 US-08-205-657A-1	Sequence 1, Appli
24	56.2	3.4	1888	4 US-08-702-525-1	Sequence 1, Appli
25	56.2	3.4	1888	5 PCT-US95-02576-1	Sequence 1, Appli
26	56.2	3.4	2516	4 US-08-205-697A-3	Sequence 3, Appli
27	56.2	3.4	2516	4 US-08-205-697A-3	Sequence 3, Appli

28	56.2	3.4	2516	4 US-08-702-525-3	Sequence 3, Appli
29	56.2	3.4	2516	5 PCT-US95-02576-3	Sequence 3, Appli
30	55.6	3.3	1906	4 US-09-031-962D-5	Sequence 5, Appli
31	55.6	3.3	80595	4 US-09-078-294-3	Sequence 3, Appli
32	55.4	3.3	965	3 US-09-280-799-133	Sequence 133, App
33	55.4	3.3	3925	4 US-08-793-044-1	Sequence 1, Appli
34	55.4	3.3	8387	2 US-08-532-814-1	Sequence 1, Appli
35	55.4	3.3	8388	4 US-09-225-509-1	Sequence 1, Appli
36	55.2	3.3	3172	1 US-07-741-940-3	Sequence 3, Appli
37	55.2	3.3	3172	1 US-08-289-548A-3	Sequence 3, Appli
38	55.2	3.3	3172	1 US-08-452-654-3	Sequence 3, Appli
39	55.2	3.3	3172	1 US-08-452-655B-3	Sequence 3, Appli
40	55.2	3.3	3172	1 US-08-450-582-3	Sequence 3, Appli
41	54.8	3.3	4157	2 US-08-871-266B-1	Sequence 1, Appli
42	54.8	3.3	4157	2 US-08-819-458A-1	Sequence 1, Appli
43	54.8	3.3	4157	2 US-09-018-864A-1	Sequence 1, Appli
44	54.8	3.3	4157	3 US-08-871-267B-1	Sequence 1, Appli
45	54.8	3.3	4157	4 US-09-618-419-1	Sequence 1, Appli

ALIGNMENTS

RESULT 1
US-08-142-439A-1
Sequence 1, Application US/08142439A
Patent No. 5670360
GENERAL INFORMATION:
APPLICANT: Thorens, Bernard
TITLE OF INVENTION: Receptor for the Glucagon-Like Peptide-1
NUMBER OF SEQUENCES: 9
CORRESPONDENCE ADDRESS:
ADDRESSEE: No. 56703600 No. 56703600disk of No. 5670360th America, Inc.
STREET: 405 Lexington Avenue, Suite 6400
CITY: New York
STATE: New York
COUNTRY: U.S.A.
ZIP: 10174-6201
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patent Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/142,439A
FILING DATE: 24-NOV-93
CLASSIFICATION: 530
PRIOR APPLICATION DATA:
APPLICATION NUMBER: DK 398/92
FILING DATE: 25-MAR-92
PRIOR APPLICATION DATA:
APPLICATION NUMBER: PCT/EP93/00697
FILING DATE: 23-MAR-93
ATTORNEY/AGENT INFORMATION:
NAME: Harrington, James J.
REGISTRATION NUMBER: 38,711
REFERENCE/DOCKET NUMBER: 3756.204-US
TELECOMMUNICATION INFORMATION:
TELEPHONE: 212 867 0123
TELEFAX: 212 867 0298
INFORMATION FOR SEQ ID NO: 1:
SEQUENCE CHARACTERISTICS:
LENGTH: 3066 base pairs
TYPE: nucleic acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: cDNA
HYPOTHETICAL: NO
ORIGINAL SOURCE:
ORGANISM: Rat
FEATURE:
NAME/KEY: CDS

Db 1772 ACACACACACACACACACACACACACACACACACACATTTTC 1713
QY 287 TTCCTCCCTCCACCTTCCTCC 308
1712 GTCCTGACGCCGGAACCCCGC 1691

RESULT 5

PCT-US95-09098-1/C
Sequence 1, Application PC/TUS9509098
GENERAL INFORMATION:
APPLICANT: Sato, Ryo
APPLICANT: Boynton, John E.
APPLICANT: Gillham, Nicholas W.
APPLICANT: Harris, Elizabeth H.
TITLE OF INVENTION: Porphyrin-Accumulating Type Herbicide
TITLE OF INVENTION: Resistance Gene
NUMBER OF SEQUENCES: 1
CORRESPONDENCE ADDRESS:
ADDRESSEE: Birch, Stewart, Kolasch & Birch, LLP
STREET: 810 Gatehouse Road, Suite 500 East
CITY: Falls Church
STATE: Virginia
COUNTRY: U.S.A.
ZIP: 22042
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patentin Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: PCT/US95/09098
FILING DATE: 20-JUL-1995
CLASSIFICATION:
ATTORNEY/AGENT INFORMATION:
NAME: Murphy Jr., Gerald M.
REGISTRATION NUMBER: 28,977
REFERENCE/DOCKET NUMBER: 2185-110P
TELECOMMUNICATION INFORMATION:
TELEPHONE: (703) 205-8000
TELEFAX: (703) 205-8050
TELEX: 248345
INFORMATION FOR SEQ ID NO: 1:
SEQUENCE CHARACTERISTICS:
LENGTH: 3383 base pairs
TYPE: nucleic acid
STRANDEDNESS: double
TOPOLOGY: linear
MOLECULE TYPE: DNA (genomic)
HYPOTHETICAL: NO
ANTI-SENSE: NO
ORIGINAL SOURCE:
ORGANISM: Chlamydomonas reinhardtii
SPRAIN: RS-3
PCT-US95-09098-1

Query Match 3.5%; Score 58; DB 5; Length 3383;
Best Local Similarity 81.7%; Pred. No. 4.2e-08;
Matches 67; Conservative 0; Mismatches 15; Indels 0; Gaps 0;

QY 227 AACACACACACACACACACACACACACACACACACATGTTTC 286
Db 1773 ACACACACACACACACACACACACACACACACACATTTTC 1714
QY 287 TTCCTCCCTCCACCTTCCTCC 308
1712 GTCCTGACGCCGGAACCCCGC 1692

RESULT 6
US-08-772-440-33
Sequence 33, Application US/08772440

Patent No. 6046158
GENERAL INFORMATION:
APPLICANT: Aritzumi, Kiyoshi
APPLICANT: Takashima, Akira
TITLE OF INVENTION: UNIQUE DENDRITIC CELL-ASSOCIATED C-TYPE
TITLE OF INVENTION: LECTINS, DECTIN-1 AND DECTIN-2; COMPOSITIONS AND USES
TITLE OF INVENTION: THEREOF
NUMBER OF SEQUENCES: 42
CORRESPONDENCE ADDRESS:
ADDRESSEE: Arnold, White & Durkee
STREET: P.O. Box 4433
CITY: Houston
STATE: Texas
COUNTRY: USA
ZIP: 77210
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patentin Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/772,440
FILING DATE: CONCURRENTLY HERewith
CLASSIFICATION: 435
ATTORNEY/AGENT INFORMATION:
NAME: Parker, David L.
REGISTRATION NUMBER: 32,165
REFERENCE/DOCKET NUMBER: UTXD:493
TELECOMMUNICATION INFORMATION:
TELEPHONE: 512/418-3000
TELEFAX: 512/474-7577
INFORMATION FOR SEQ ID NO: 33:
SEQUENCE CHARACTERISTICS:
LENGTH: 10409 base pairs
TYPE: nucleic acid
STRANDEDNESS: single
TOPOLOGY: linear
FEATURE:
NAME/KEY: modified_base
LOCATION: 6510
OTHER INFORMATION: /mod_base= OTHER
OTHER INFORMATION: /note= "D = A or G or T"
FEATURE:
NAME/KEY: modified_base
LOCATION: 3406..6470
OTHER INFORMATION: /mod_base= OTHER
OTHER INFORMATION: /note= "K = G or T"
FEATURE:
NAME/KEY: modified_base
LOCATION: 3564..7896
OTHER INFORMATION: /mod_base= OTHER
OTHER INFORMATION: /note= "M = A or C"
FEATURE:
NAME/KEY: modified_base
LOCATION: 3497..3607
OTHER INFORMATION: /mod_base= OTHER
OTHER INFORMATION: /note= "N = A or C or G or T"
FEATURE:
NAME/KEY: modified_base
LOCATION: 3479..6422
OTHER INFORMATION: /mod_base= OTHER
OTHER INFORMATION: /note= "R = A or G"
FEATURE:
NAME/KEY: modified_base
LOCATION: 3405..6871
OTHER INFORMATION: /mod_base= OTHER
OTHER INFORMATION: /note= "S = C or G"
FEATURE:
NAME/KEY: modified_base
LOCATION: 3457..9998
OTHER INFORMATION: /mod_base= OTHER
OTHER INFORMATION: /note= "W = A or T"
FEATURE:

NUMBER OF SEQ ID NOS: 22
SOFTWARE: PatentIn Ver. 2.0
SEQ ID NO 17
LENGTH: 48974
TYPE: DNA
ORGANISM: Mus musculus
US-08-920-422-17

Query Match 3.5%; Score 57.4; DB 4; Length 48974;
Best Local Similarity 80.7%; Pred. No. 3.8e-07;
Matches 67; Conservative 0; Mismatches 16; Indels 0; Gaps 0;

QY 197 ATGGACAGATACCTTACAGCTACATTCACAAAACACACACACACACACACACAC 256
DB 35831 AAGTAGAAGATAGCTGTCTATTAAGCTACACACACACACACACACACACAC 35890
QY 257 ACACACACACACACACACACAC 279
DB 35891 ACACACACACACACACACACAC 35913

RESULT 9
US-08-750-064-1
Sequence 1, Application US/08750064
Patent No. 6040142

GENERAL INFORMATION:
APPLICANT: MELKI, JUDITH
APPLICANT: MUNNICH, ARNOLD
TITLE OF INVENTION: METHOD AND PROBES FOR DETECTING MARKERS
TITLE OF INVENTION: BOUND TO THE LOCUS OF CHILD SPINAL MUSCULAR ATROPHIES
NUMBER OF SEQUENCES: 23
CORRESPONDENCE ADDRESS:
ADDRESSEE: NIXON & VANDERHAYE P.C.
STREET: 1100 NORTH GLEBE ROAD
CITY: ARLINGTON
STATE: VIRGINIA
COUNTRY: U.S.A.
ZIP: 22201-4714

COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: PatentIn Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/750,064
FILING DATE: 29-JAN-1997
CLASSIFICATION: 435
PRIOR APPLICATION DATA:
APPLICATION NUMBER: FR 94/06856
FILING DATE: 03-JUN-1994
ATTORNEY/AGENT INFORMATION:
NAME: WILSON, MARY J.
REGISTRATION NUMBER: 32,955
REFERENCE/DOCKET NUMBER: 960-26
TELECOMMUNICATION INFORMATION:
TELEPHONE: (703) 816-4000
TELEFAX: (703) 816-4100
INFORMATION FOR SEQ ID NO: 1:
SEQUENCE CHARACTERISTICS:
LENGTH: 372 base pairs
TYPE: nucleotide
STRANDEDNESS: double
TOPOLOGY: linear
MOLECULE TYPE: DNA (genomic)
HYPOTHETICAL: NO
ANTI-SENSE: NO
ORIGINAL SOURCE:
ORGANISM: Homo sapiens
US-06-750-064-1

Query Match 3.4%; Score 57.2; DB 3; Length 372;
Best Local Similarity 75.5%; Pred. No. 1.8e-08;
Matches 71; Conservative 0; Mismatches 23; Indels 0; Gaps 0;

QY 224 TCAAAACACACACACACACACACACACACACACACACACACATGTT 283
DB 85 TCAAAACATACACACACACACACACACACACACACACACACACACTCTCTCT 144
QY 284 TTCTTCCCTCCCTCCACATTTCTCCGATTCCTCTGT 317
DB 145 CTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 178

RESULT 10
US-08-545-196B-14
Sequence 14, Application US/08545196B
Patent No. 6080577

GENERAL INFORMATION:
APPLICANT: MELKI, JUDITH
APPLICANT: MUNNICH, ARNOLD
TITLE OF INVENTION: SURVIVAL MOTOR NEURON (SMN) GENE: A GENE
TITLE OF INVENTION: FOR SPINAL MUSCULAR ATROPHY
NUMBER OF SEQUENCES: 65
CORRESPONDENCE ADDRESS:
ADDRESSEE: BIRCH, STEWART, KOLASCH AND BIRCH, LLP
STREET: PO BOX 747
CITY: FALLS CHURCH
STATE: VA
COUNTRY: USA
ZIP: 22040-0747

COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: PatentIn Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/545,196B
FILING DATE: 19-OCT-1995
CLASSIFICATION: 435
ATTORNEY/AGENT INFORMATION:
NAME: FARACI, C. J.
REGISTRATION NUMBER: 32,350
REFERENCE/DOCKET NUMBER: 2121-110P
TELECOMMUNICATION INFORMATION:
TELEPHONE: (703) 205-8000
TELEFAX: (703) 205-8050
INFORMATION FOR SEQ ID NO: 14:
SEQUENCE CHARACTERISTICS:
LENGTH: 372 base pairs
TYPE: nucleic acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: other nucleic acid
DESCRIPTION: /desc = "SYNTHETIC DNA"

US-08-545-196B-14

Query Match 3.4%; Score 57.2; DB 3; Length 372;
Best Local Similarity 75.5%; Pred. No. 1.8e-08;
Matches 71; Conservative 0; Mismatches 23; Indels 0; Gaps 0;

QY 224 TCAAAACACACACACACACACACACACACACACACACACACATGTT 283
DB 85 TCAAAACATACACACACACACACACACACACACACACACACACACTCTCTCT 144
QY 284 TTCTTCCCTCCCTCCACATTTCTCCGATTCCTCTGT 317
DB 145 CTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 178

RESULT 11
US-09-078-294-4
Sequence 4, Application US/09078294
Patent No. 6265211
GENERAL INFORMATION:
APPLICANT: Choo, Kong-Hong Andy
APPLICANT: Du Sart, Desirée

TOPOLOGY: linear
IMMEDIATE SOURCE:
CLONE: pTZ9pt-F1s
US-08-232-463-14

Query Match 3.4%; Score 56.4; DB 1; Length 7218;
Best Local Similarity 10.3%; Pred. No. 2.3e-07;
Matches 42; Conservative 194; Mismatches 170; Indels 0; Gaps 0;

```

OY 385 TTCGAGACTGATCTTTTATCATTAAGGTTTGAATCTTGCCAGCTGGGTTTAAAGT 444
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB 1531 TTCTAGACCACTATCTACATTTCAAAAACGCGATGAGCATCTGTAATTACCTAT 1472
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
OY 445 TTTAGGATTTTATCTAGCGCATCTCCTCCCTGTAATGTCAGAACTCACT 504
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB 1471 CTATCGAGTATGTTAAAGAGATAGAAGATTGGTACRRRRRRRRRRRRRRRRR 1412
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
OY 505 GGGCTTGCTACCTAATGGAATGATCTATGGTTGACTTAATGTGAAGAAAAA 564
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB 1411 RRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRR 1352
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
OY 565 GAAGGGGAAAAGAGGAGGAGGAGAAAGAGGAAAGAACTGCTTTATGCCATT 624
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB 1351 RRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRR 1292
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
OY 625 GCTACTTAACATTTTGTCTCACCCTCCACTGTTCTTCAATGGAAGACTGATAG 684
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB 1291 RRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRR 1232
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
OY 685 AAGCTGGAGCCAGCCAGGATAGAGAGTGTGTGTGTGGGGGGGTGGGAG 744
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB 1231 RRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRR 1172
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
OY 745 CAAGCAGACCTTAGACAGAGAGAGCCTGTAGATGATGAG 790
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB 1171 RRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRR 1126
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||

```

Search completed: November 6, 2002, 14:19:56
Job time : 417.027 secs

GenCore version 5.1.3
Copyright (c) 1993 - 2002 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: November 6, 2002, 14:03:45 ; Search time 259.209 Seconds
(without alignments)
10995.308 Million cell updates/sec

Title: US-09-833-740-1_COPY_1_1660

Perfect score: 1660

Sequence: 1 aggttaccgacttgcaca.....gaagagactgtgcacaaca 1660

Scoring table: IDENTITY NUC

Gapop 10.0 , Gapext 1.0

Searched: 1736436 seqs, 858457221 residues

Total number of hits satisfying chosen parameters: 3472872

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

N.Geneseq_032802.*
1: /SIDSI/gcgdata/geneseq/geneseqn-emb1/NA1980.DAT.*
2: /SIDSI/gcgdata/geneseq/geneseqn-emb1/NA1981.DAT.*
3: /SIDSI/gcgdata/geneseq/geneseqn-emb1/NA1982.DAT.*
4: /SIDSI/gcgdata/geneseq/geneseqn-emb1/NA1983.DAT.*
5: /SIDSI/gcgdata/geneseq/geneseqn-emb1/NA1984.DAT.*
6: /SIDSI/gcgdata/geneseq/geneseqn-emb1/NA1985.DAT.*
7: /SIDSI/gcgdata/geneseq/geneseqn-emb1/NA1986.DAT.*
8: /SIDSI/gcgdata/geneseq/geneseqn-emb1/NA1987.DAT.*
9: /SIDSI/gcgdata/geneseq/geneseqn-emb1/NA1988.DAT.*
10: /SIDSI/gcgdata/geneseq/geneseqn-emb1/NA1989.DAT.*
11: /SIDSI/gcgdata/geneseq/geneseqn-emb1/NA1990.DAT.*
12: /SIDSI/gcgdata/geneseq/geneseqn-emb1/NA1991.DAT.*
13: /SIDSI/gcgdata/geneseq/geneseqn-emb1/NA1992.DAT.*
14: /SIDSI/gcgdata/geneseq/geneseqn-emb1/NA1993.DAT.*
15: /SIDSI/gcgdata/geneseq/geneseqn-emb1/NA1994.DAT.*
16: /SIDSI/gcgdata/geneseq/geneseqn-emb1/NA1995.DAT.*
17: /SIDSI/gcgdata/geneseq/geneseqn-emb1/NA1996.DAT.*
18: /SIDSI/gcgdata/geneseq/geneseqn-emb1/NA1997.DAT.*
19: /SIDSI/gcgdata/geneseq/geneseqn-emb1/NA1998.DAT.*
20: /SIDSI/gcgdata/geneseq/geneseqn-emb1/NA1999.DAT.*
21: /SIDSI/gcgdata/geneseq/geneseqn-emb1/NA2000.DAT.*
22: /SIDSI/gcgdata/geneseq/geneseqn-emb1/NA2001A.DAT.*
23: /SIDSI/gcgdata/geneseq/geneseqn-emb1/NA2001B.DAT.*
24: /SIDSI/gcgdata/geneseq/geneseqn-emb1/NA2002.DAT.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	1658.8	99.9	2170	AAH43758	5' flanking and 5'
2	64.6	3.9	1199	AAQ36051	Human nov gene pro
3	63.4	3.8	245	AAV30918	Human secreted pro
4	63.4	3.8	245	AAF98397	3' portion of cDNA
5	61.8	3.7	353	AAI88828	Human polynucleoti
6	61.2	3.7	3809	AAO12760	P40 genomic DNA.
7	60.6	3.7	6183	ABL33197	Human immune syste
8	60.6	3.7	23885	AAK70103	Human immune/haema
9	60.6	3.7	23885	AAK73173	Human immune/haema

10	60.4	3.6	415	AAI90376	Human polynucleoti
11	60.4	3.6	8095	AAK81626	Human immune/haema
12	60.2	3.6	4858	AAQ91904	Murine A259 coding
13	60.2	3.6	4858	AAI6874	Murine A259 cDNA.
14	60	3.6	1787	AAA36622	Murine stomatin DN
15	59.8	3.6	5176	AAK66814	Human immune/haema
16	59.8	3.6	56153	AAK46793	Tumour suppressor
17	59.6	3.6	19062	AAK21280	Human low adenosin
18	59.6	3.6	22142	AAA35158	Human adenosine re
19	59.6	3.6	30417	AAK21282	Human low adenosin
20	59.6	3.6	30417	AAA35160	Human adenosine re
21	59.2	3.6	5275	AAK46378	Human adenosine re
22	59.2	3.6	5275	ABL32825	Tumour suppressor
23	59.2	3.6	160271	AAK85750	Human immune syste
24	59.2	3.6	160271	AAK85756	Bipolar affective
25	59.2	3.6	160271	AAK85756	Human chromosome 1
26	59.2	3.6	160271	AAK85756	Human chromosome 1
27	59.2	3.6	160271	AAK85756	Human chromosome 1
28	59.2	3.6	160271	AAK85756	Human chromosome 1
29	59.2	3.6	160271	AAK85756	Human chromosome 1
30	59.2	3.6	160271	AAK85756	Human chromosome 1
31	59	3.6	54548	AAK85756	Nucleotide sequenc
32	58.8	3.5	3304	AAK25360	DNA sequence of th
33	58.8	3.5	5165	ABL26068	NSQO gene-12 assoc
34	58.8	3.5	168575	AAK21613	Drosophila melanog
35	58.6	3.5	143068	AAK21105	Human hypochlorin r
36	58.6	3.5	143068	AAK21105	Human low adenosin
37	58.6	3.5	143068	AAK21272	Human low adenosin
38	58.6	3.5	143068	AAK4983	Human adenosine re
39	58.6	3.5	149412	AAA35150	Human adenosine re
40	58.6	3.5	152740	AAK21273	Human adenosine re
41	58.4	3.5	405	AAK21273	Human low adenosin
42	58.4	3.5	437	AAI88261	Human polynucleoti
43	58.4	3.5	574	ABA59754	Human polynucleoti
44	58.4	3.5	574	ABA59754	Human foetal liver
45	58.4	3.5	574	AAK08024	Probe #6724 for ge
					Human brain expres

ALIGNMENTS

RESULT 1	AAH43758	AAH43758 standard; DNA; 2170 BP.
ID	AAH43758	
XX	AAH43758;	
AC	AAH43758;	
XX		
DT	30-JAN-2002	(first entry)
XX		
DE	5' flanking and 5' UTR of GLP-2 receptor gene.	
XX		
KW	5' flanking region; 5' UTR; untranslated region; murine;	
KW	glucagon-like peptide-2; GLP-2; receptor; regulator;	
KW	intestinal epithelium; lateral hypothalamus; promoter; ds.	
XX		
OS	Mus musculus.	
XX		
FH	Key	Location/Qualifiers
FT	misc_RNA	1656
FT		/tag= a
FT		/note= "Punitive transcriptional start site"
FT	misc_RNA	1761..1763
FT		/tag= b
FT		/note= "Corresponds to translational start in rat/human
FT		GLP-2R gene"
FT	CDS	1875..1877
FT		/tag= c
FT		/note= "Putative translational start site"
PN	W0200179290-A2.	
XX		
XX		
XX		
PD	25-OCT-2001.	
XX		

PF 12-APR-2001; 2001MO-IB00619.
XX
PR 13-APR-2000; 2000US-196909P.
PR 01-FEB-2001; 2001US-265310P.
XX
PA (ONEO-) 1149336 ONTARIO INC.
XX
PI Drucker DJ, Lovshin JAL;
XX
XX WPI; 2002-026012/03.
XX
PT Novel recombinant DNA construct, useful as drug screening tool and as
PT protein delivery vehicle, comprising glucagon-like peptide-2 receptor
PT gene, and linked to heterologous gene of interest for expression -
XX
XX Disclosure; Fig 1; 70pp; English.

This sequence represents the 5' flanking and 5' UTR of the murine
CC glucagon-like peptide-2 (GLP-2) receptor gene. This sequence may be
CC used in the DNA construct of the invention, such that it is linked
CC for expression with a heterologous gene of interest. This construct
CC is useful for screening compounds to identify regulators of GLP-2
CC receptor expression. The DNA construct is also useful for delivering
CC a gene product to tissues selected from intestinal epithelium and the
CC lateral hypothalamus, by transfecting an organism, or a gametic or
CC embryonic form of such an organism, with the construct. The DNA
CC construct is also useful to identify cells capable of mediating
CC expression from the chosen GLP-2 promoter, to identify regions of
CC the GLP-2 promoter that are functional in a given cell type, and
CC to screen for agents that modulate expression from the GLP-2 promoter.
CC It can be used to map functional regions of the GLP-2 promoter, and
CC for screening agents that modulate the function of GLP-2 promoter.
XX

Sequence 2170 BP; 574 A; 501 C; 550 G; 540 T; 5 other:

Query Match 99.9%; Score 1658.8; DB 24; Length 2170;
Best Local Similarity 100.0%; Pred. No. 0;
Matches 1660; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 AGCTTTACCGGATTTTACACTAGATGCGATCCGCGGAGGTAGCGATGAAGCT 60
DB 1 AGCTTTACCGGATTTTACACTAGATGCGATCCGCGGAGGTAGCGATGAAGCT 60
QY 61 GACCAAGCAAGCTCTTACAGGGGAGAAATCGAGAAAAAATTTTGAAGCAT 120
DB 61 GACCAAGCAAGCTCTTACAGGGGAGAAATCGAGAAAAAATTTTGAAGCAT 120
QY 121 TTCAAGAAAGCAAGATGCAATATTTGTACAAAACAGTCTTCTCCACCATGCGAC 180
DB 121 TTCAAGAAAGCAAGATGCAATATTTGTACAAAACAGTCTTCTCCACCATGCGAC 180
QY 181 CGGAGAGCTCCGATATGACAGATAGCTTTTACAGTACATTCAAAACACACACAC 240
DB 181 CGGAGAGCTCCGATATGACAGATAGCTTTTACAGTACATTCAAAACACACACAC 240
QY 241 ACACACACACACACACACACACACACACACACATGTTTCTCCCTCCCTCAC 300
DB 241 ACACACACACACACACACACACACACACACACATGTTTCTCCCTCCCTCAC 300
QY 301 TTCTCTCCATTTCTGTGGTCCCAAGAGATGACATATTTAGCTGTAGAAATCACAC 360
DB 301 TTCTCTCCATTTCTGTGGTCCCAAGAGATGACATATTTAGCTGTAGAAATCACAC 360
QY 361 CATTAAGAGCCATCGGAGGACATTTCCAGACTGATCTTTATCATTAAGGTTGAAT 420
DB 361 CATTAAGAGCCATCGGAGGACATTTCCAGACTGATCTTTATCATTAAGGTTGAAT 420
QY 421 CTTCGCAAGTGTGGATTTTAAGGATTTTATCTAGCGGCACTGACCTGCTT 480
DB 421 CTTCGCAAGTGTGGATTTTAAGGATTTTATCTAGCGGCACTGACCTGCTT 480
QY 481 CCTGTGATGTTTCAAGATTCAGTGGCTTGTGCTAGCTAATGGAATGCTATGTTTG 540
DB 481 CCTGTGATGTTTCAAGATTCAGTGGCTTGTGCTAGCTAATGGAATGCTATGTTTG 540

DB 481 CCTGTGATGTTTCAAGATTCAGTGGCTTGTGCTAGCTAATGGAATGCTATGTTTG 540
QY 541 ACTTAATGTAAGGAAAAAAGAGGGGAGAAAGAGGAGGAGGAGGAGGAG 600
DB 541 ACTTAATGTAAGGAAAAAAGAGGGGAGAAAGAGGAGGAGGAGGAGGAGGAG 600
QY 601 GGAAGAGCTGCTTTTATGCTATTTGCTACTTAACATTTTGTCTCACCCTTCAGCTTGG 660
DB 601 GGAAGAGCTGCTTTTATGCTATTTGCTACTTAACATTTTGTCTCACCCTTCAGCTTGG 660
QY 661 TTCTTCAATGGAAGAGCTGATAGAAAGCTGGAGGCCAGGAGGATAGGAGAGTGTGT 720
DB 661 TTCTTCAATGGAAGAGCTGATAGAAAGCTGGAGGCCAGGAGGATAGGAGAGTGTGT 720
QY 721 GTGTGTGTGGGGGGGGGGGGGGGAGCAAGAGAGCTTATAGACAGAGAGAGAGCTGTAG 780
DB 721 GTGTGTGTGGGGGGGGGGGGGGGAGCAAGAGAGCTTATAGACAGAGAGAGAGCTGTAG 780
QY 781 AGAYCATGAGCTTCTTTGAGACCCCTAGTGTAAACAGAGATAGTCTTAACAGGTAGC 840
DB 781 AGAYCATGAGCTTCTTTGAGACCCCTAGTGTAAACAGAGATAGTCTTAACAGGTAGC 840
QY 841 TGTGTGACGTGACTGCTGCTGGAAGSCTGTGCTTTTCTGCTGCTGTGAGCTT 900
DB 841 TGTGTGACGTGACTGCTGCTGGAAGSCTGTGCTTTTCTGCTGCTGTGAGCTT 900
QY 901 GAACAAACACCCGCGCTCTTTGAAACCCCACTATTTCTCAGCCCTCAGATGAAGATTA 960
DB 901 GAACAAACACCCGCGCTCTTTGAAACCCCACTATTTCTCAGCCCTCAGATGAAGATTA 960
QY 961 TGTGTGCTGGAAGATAGTATGAGGTTCAAGTGAAGTGAAGGAGGAGGAGGAGGTTTG 1020
DB 961 TGTGTGCTGGAAGATAGTATGAGGTTCAAGTGAAGTGAAGGAGGAGGAGGAGGTTTG 1020
QY 1021 TAACCATTAACCTAAGTGGGGGTGTGTTAGTATGATGATGATGATGATGATGATGATG 1080
DB 1021 TAACCATTAACCTAAGTGGGGGTGTGTTAGTATGATGATGATGATGATGATGATGATG 1080
QY 1081 TGTGAGTGGCCAGAGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 1140
DB 1081 TGTGAGTGGCCAGAGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 1140
QY 1141 CAATTGAAAGGACCAACCTGCTGCTTCTTCCATCCGTTGTTTGTACCGGTAAACGCA 1200
DB 1141 CAATTGAAAGGACCAACCTGCTGCTTCTTCCATCCGTTGTTTGTACCGGTAAACGCA 1200
QY 1201 ACTCACTCTGCGGACAGCAAGGCTTTTGTGAGAGCTGAGAGCTGAGTGTCT 1260
DB 1201 ACTCACTCTGCGGACAGCAAGGCTTTTGTGAGAGCTGAGAGCTGAGTGTCT 1260
QY 1261 CTGAGACAGGAAACATCATCTTGTACTATGAGCATAGTATGATGATGATGATGATGAT 1320
DB 1261 CTGAGACAGGAAACATCATCTTGTACTATGAGCATAGTATGATGATGATGATGATGAT 1320
QY 1321 AGCCCTGAGCTGCTGCTGCTTGAAGAAAGTTGATGCTTTTGAAGAAAGAAATCTGTGCTTAA 1380
DB 1321 AGCCCTGAGCTGCTGCTGCTTGAAGAAAGTTGATGCTTTTGAAGAAAGAAATCTGTGCTTAA 1380
QY 1381 AGAAGCTTACTGAGCATGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 1440
DB 1381 AGAAGCTTACTGAGCATGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 1440
QY 1441 GTGCGTAAAGATAGATCTGCAATGATGATGATGATGATGATGATGATGATGATGATGAT 1500
DB 1441 GTGCGTAAAGATAGATCTGCAATGATGATGATGATGATGATGATGATGATGATGATGAT 1500
QY 1501 TGAAGAAAGTCCAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 1560
DB 1501 TGAAGAAAGTCCAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 1560
QY 1561 TTTTGAAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 1620
DB 1561 TTTTGAAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 1620

OY 1621 ACTCAGACAGTCGGCGCCTGAAGAGACTTGTGCAACA 1660
|||||
DB 1621 ACTCAGACAGTCGGCGCCTGAAGAGACTTGTGCAACA 1660

RESULT 2

AAQ36051/c
ID AAQ36051 standard; cDNA: 1199 BP.

AC AAQ36051;

DT 24-MAY-1993 (first entry)

XX Human nov gene promoter sequence XXXIII fragment 2.

XX avian nephroblastoma; avian myeloblastoma virus;

KW stringent hybridisation; ss.

OS Homo sapiens.

XX WO9300430-A.

XX 07-JAN-1993.

XX 25-JUN-1992: 92MO-FR00589.

XX 25-JUN-1991: 91FR-0007807.

PA (CNRS) CENT NAT RECH SCI.

XX Martinerie C, Perbal B;

DR WPI: 1993-036377/04.

PT Nucleotide sequences hybridising to regions of chicken nov gene -

PT useful as probes for detecting complementary sequences to

XX evaluate development and/or differentiation of tumours

PS Disclosure; Page 45-46; 67pp; French.

CC The invention includes nucleotide sequences which can hybridise to

CC all or part of the chicken nov gene under stringent conditions

CC (i.e. 50% formamide, 5 x SSC). The promoter sequence from the human

CC nov gene, localised to a 2.2kb PstI-HindII fragment and comprising

CC the 283bp upstream of the start of exon 1 is specifically

CC disclosed. The promoter sequence corresponds to sequence XXXIII

CC which is given as three separate fragments in the specification.

CC The human nov promoter contains several consensus sequences for

CC different transcription factors such as NF1, AP1 and Sp1. It also

CC comprises a sequence of 20 repeats of the TG motif (a possible

CC polymorphic marker). See also AAQ36050 and AAQ36052.

XX Sequence 1199 BP; 269 A; 292 C; 248 G; 390 T; 0 other;

XX Query Match 3.9%; Score 64.6; DB 14; Length 1199;

XX Best Local Similarity 57.8%; Pred. No. 2.6e-08;

XX Matches 115; Conservative 0; Mismatches 84; Indels 0; Gaps 0;

DB 845 ACACACACACATTAAGAAA 827

RESULT 3

AAV30918
ID AAV30918 standard; DNA: 245 BP.

XX AAV30918;

XX 14-SEP-1998 (first entry)

XX Human secreted protein AS63_29 cDNA 3' region.

XX AS63_29; secreted protein; protein factor; human; ds.

XX Homo sapiens.

XX WO9817687-A2.

XX 30-APR-1998.

XX 24-OCT-1997: 97MO-US19590.

XX 24-OCT-1997: 97US-0740274.

XX 25-OCT-1996: 96US-0740274.

PA (GENM) GENETICS INST INC.

XX Agostino MJ, Jacobs K, Lavallie ER, McCoy JM, Merberg D;

PI Racie LA, Spaulding V, Treacy M;

DR WPI: 1998-261426/23.

XX Nucleic acid encoding secreted protein from human cells - useful,

PT e.g. as immunomodulator, antitumour agent, promoters of tissue

PT growth, haemostatic and thrombolytic agents etc.

XX Claim 16; Page 69; 114pp; English.

CC This polynucleotide comprises the 3' portion of a cDNA clone,

CC designated AS63_29 (see also AAV30917), that codes for a novel human

CC secreted protein (see AAV58381). AS63_29 was isolated from a human

CC foetal brain cDNA library using methods selective for cDNAs that

CC encode secreted proteins. The clone is deposited in composite clone

CC ATCC 98232: an oligonucleotide (see AAT99723) is designed to isolate

CC the clone from the composite. Novel cDNA clones (see AAV30916-32)

CC coding for human secreted proteins (see AAV58580-90) are claimed.

CC These can be used for recombinant production of the secreted

CC proteins for analysis, characterisation, diagnostic or therapeutic

XX use.

XX Sequence 245 BP; 85 A; 69 C; 38 G; 45 T; 8 other;

XX Query Match 3.8%; Score 63.4; DB 19; Length 245;

XX Best Local Similarity 77.6%; Pred. No. 2.4e-08;

XX Matches 76; Conservative 0; Mismatches 22; Indels 0; Gaps 0;

DB 133 CTTCGCTCCNTCCCGCAGTGGCTTGTGTGATCGCAAG 170

RESULT 4

AAF98397
ID AAF98397 standard; cDNA: 245 BP.

XX AAF98397;

XX 07-JUN-2001 (first entry)

XX 3' portion of cDNA encoding AS63_29 protein SEQ ID 38.
DE Human; secreted protein; nutrient; cytokine modulator; proliferation;
XX differentiation; immune system modulator; tissue growth; chemotactic;
KW haemostatic; thrombolytic; anti-inflammatory; tumour inhibition; ss;
KW haematopoiesis.
XX Homo sapiens.
OS MO200119988-A1.
XX
XX
XX 22-MAR-2001.
XX
XX 14-SEP-2000; 2000MO-US25135.
XX
XX 17-SEP-1999; 99US-0398829.
XX
XX (GENY) GENETICS INST INC.
XX
XX Jacobs K, McCoy JM, Lavallie ER, Collins-Racie LA, Evans C;
PI Merberg D, Treacy M, Bowman MR, Spaulding V, Agostino MJ;
XX WPI; 2001-244801/25.
XX
XX Isolated nucleic acids encoding polypeptides, useful for modulating
PT e.g. cytokine and cell proliferation/differentiation activity, the
PT immune system and hematopoiesis regulating activity -
XX
XX Disclosure; Page 405; 557bp; English.
XX
XX Human cDNA clones represented in AAF98374 - AAF98489 encode secreted
CC proteins AAB90667 - AAB90750. The cDNA clones are isolated from various
CC tissue types, and may be used in the prevention, treatment and diagnosis
CC of diseases associated with inappropriate protein expression. The
CC polypeptides and nucleic acids may be used as nutrients or to modulate
CC cytokine and cell proliferation/differentiation activity and may also be
CC involved in modulation of the immune system. The cDNA sequences,
CC proteins, their agonists and/or antagonists exhibit haematopoiesis
CC regulating activity; tissue growth activity; activin/inhibin activity;
CC chemotactic/chemokinetic activity; haemostatic and thrombolytic
CC activity; receptor/ligand activity; anti-inflammatory activity;
CC haematopoiesis activity; cadherin/tumour suppressor activity; and/or
CC tumour inhibition activity. Included in the invention are probes
CC represented in AAF98490 - AAF98572 which are specific for the cDNA clones
CC encoding the secreted proteins.
XX
XX Sequence 245 BP; 85 A; 69 C; 38 G; 45 T; 8 other;
SQ
Query Match 3.8%; Score 63.4; DB 22; Length 245;
Best Local Similarity 77.6%; Pred. No. 2.4e-08;
Matches 76; Conservative 0; Mismatches 22; Indels 0; Gaps 0;
OY 230 CATGTTTCTC 289
DB 73 CATTTTGATCC 132
OY 290 CCTCCCTCCACTTCCTCCATTCCTGTGTGTCCTCAAG 327
DB 133 CTTGCTTCCTCCCGCCGATGCTGTGTGATCGCAAG 170
RESULT 5
AAI88828/c
ID AAI88828 standard; cDNA: 353 BP.
XX
XX AAI88828;
AC
XX
XX 06-NOV-2001 (first entry)
DT
XX
XX Human polynucleotide SEQ ID NO 8888.
DE
XX
XX Human; cytokine; cell proliferation; cell differentiation; gene therapy;

KW vaccine; peptide therapy; stem cell growth factor; haematopoiesis;
KW tissue growth factor; immunomodulatory; cancer; leukaemia;
KW nervous system disorders; arthritis; inflammation; ss.
XX
XX Homo sapiens.
OS MO200164835-A2.
XX
XX
XX 07-SEP-2001.
XX
XX 26-FEB-2001; 2001MO-US04927.
XX
XX 28-FEB-2000; 2000US-0515126.
XX 18-MAY-2000; 2000US-0577409.
XX
XX (HYSE-) HYSEQ INC.
XX
XX Tang YT, Liu C, Drmanac RT;
PI WPI; 2001-514838/56.
XX P-PSDB; AAO08897.
XX
XX Claim 1; SEQ ID NO 8888; 1399pp + Sequence Listing; English.
XX
XX The invention relates to human polynucleotides (AAI79941-AAI93841) and
CC the encoded proteins (AAO00010-AAO13910) that exhibit activity relating to
CC cytokine, cell proliferation or cell differentiation or which may induce
CC production of other cytokines in other cell populations. The
CC polynucleotides and polypeptides are useful in gene therapy, vaccines or
CC peptide therapy. The polypeptides have various cytokine-like activities,
CC e.g. stem cell growth factor activity, haematopoiesis regulating
CC activity, tissue growth factor activity, immunomodulatory activity and
CC activin/inhibin activity and may be useful in the diagnosis and/or
CC treatment of cancer, leukaemia, nervous system disorders, arthritis and
CC inflammation.
CC Note: The sequence data for this patent did not form part of the printed
CC specification, but was obtained in electronic format directly from WIPO
CC at ftp.wipo.int/pub/published_pcl_sequences.
XX
XX Sequence 353 BP; 65 A; 59 C; 108 G; 121 T; 0 other;
SQ
Query Match 3.7%; Score 61.8; DB 22; Length 353;
Best Local Similarity 80.9%; Pred. No. 8.7e-08;
Matches 72; Conservative 0; Mismatches 17; Indels 0; Gaps 0;
OY 226 AAAACATGTTT 285
DB 210 AAAACAGGCAAGT 151
OY 286 CTTCCCTCCCTCCACTTCCTCCATTCCTC 314
DB 150 ATTCCCTCCAGCCGATTCCTCCACATCTC 122
RESULT 6
AAQ12760
ID AAQ12760 standard; DNA: 3809 BP.
XX
XX AAQ12760;
AC
XX
XX 17-OCT-1991 (first entry)
DT
XX
XX P40 genomic DNA.
DE
XX
XX T cell growth factor; AIDS; cytokine; T helper cell; ss.
KW
XX
XX Mus musculus.
OS
XX
XX key Location/Qualifiers
FH

FT	exon	644..757
FT		/tag= a
FT		/number= 1
FT	exon	846..881
FT		/tag= b
FT		/number= 2
FT	exon	965..997
FT		/tag= c
FT		/number= 3
FT	exon	2184..2315
FT		/tag= d
FT		/number= 4
FT	exon	3393..3512
FT		/tag= e
FT		/number= 5
FT	misc_signal	227..233
FT		/tag= f
FT		/label= C-rich
FT	misc_signal	251..255
FT		/tag= g
FT		/label= C-rich
FT	misc_signal	475..480
FT		/tag= h
FT		/note= "interferon regulatory factor-1 binding element"
FT	misc_signal	478..484
FT		/tag= i
FT		/note= "AP-1 binding site"
FT	TATA_signal	589..594
FT		/tag= j
FT	sig_peptide	644..697
FT		/tag= k
FT	misc_RNA	1361..1407
FT		/tag= l
FT		/label= polyd-AdC_region
FT		/note= "% DNA; enhancer activity"
FT	polya_signal	3582..3588
FT		/tag= m
FT	polya_signal	3678..3683
FT		/tag= n
FT	misc_RNA	3523..3529
FT		/tag= on
FT		/label= ATTTTA_motif
FT		/note= "reduces stability of mRNA"
FT	misc_RNA	3534..3538
FT		/tag= p
FT		/label= ATTTTA_motif
FT		/note= "reduces stability of mRNA"
FT	misc_RNA	3543..3547
FT		/tag= q
FT		/label= ATTTTA_motif
FT		/note= "reduces stability of mRNA"
PN	WO9110738-A.	
PN		
PD	25-JUL-1991.	
XX		
XX	08-JAN-1991;	91WO-US00145.
XX		
XX	08-JAN-1990;	91WO-US00145.
XX		
XX	08-JAN-1990;	90US-0462158.
XX		
PA	(RESE) RES CORP TECHN INC.	
XX		
XX	Vansnick J, Uytenhove C, Simpson RJ;	
PI		
DR	WPI: 1991-238026/32.	
DR	P-PSDB: AARI3218.	
XX		
XX	Nucleic acid encoding for P40 T-cell growth factor - for patie	
PT	with AIDS or compromised immune systems, also allows increased	
PT	prodn. of other cytokine(s).	
XX		
XX	Claim 1; Fig 15; 122pp; English.	

CC	The sequence was obtd. from a clone, lambda S40.1a, isolated from
CC	a genomic library using a murine P40 cDNA clone as a probe. The
CC	sequence, or the cDNA sequence obtd. using the sequence, can be
CC	inserted into a vector for expression of P40 in a host organism.
CC	The protein is useful for stimulating the proliferation of certain
CC	subsetts of T helper cells e.g. in AIDS patients or immune
CC	compromised patients.
CC	See also AAQ12759 and AAQ13660-Q13663.
XX	
S0	Sequence 3809 BP; 1013 A; 903 C; 869 G; 1024 T; 0 other;
XX	
Query Match	3.7%; Score 61.2; DB 12; Length 3809;
Best Local Similarity	76.5%; Pred. No. 4.9e-07;
Matches 75; Conservative	0; Mismatches 23; Indels 0; Gaps 0
OY	220 ACATTCAAAACAGACACACACACACACACACACACACACACACACACA 279
Db	1352 ATATGACAAACAGACACACACACACACACACACACACACACACACACC 1411
OY	280 TGTTTCTTCCTCCTCCACTTCTCTCCATTCGTGT 317
Db	1412 CTTCCTCGCATCTCTACTCTTCCCTCCAGT 1449
RESULT 7	
ABL3J197/c	
ID	ABL3J197 standard; DNA: 6183 BP.
XX	
AC	ABL3J197;
XX	
DT	26-MAR-2002 (first entry)
XX	
DE	Human immune system associated gene SEQ ID NO: 1170.
XX	
KW	Human: immune system disease; cytosine methylation; antiasthmatic;
KM	antiatleriosclerotic; antianaemic; cytosolic; noctropic;
KM	neuroprotective; anti-HIV; anticoagulant; ophthalmological;
KM	antiinfective; antiarthritic; antididiabetic; antiporiatic;
KW	antiinflammatory; cancer; eye disease; arteriosclerosis; anaemia;
KV	acute myeloid leukaemia; Alzheimer's disease; AIDS; epilepsy;
KM	neurofibromatosis; rheumatoid arthritis; psoriasis; bowel disease;
KX	gene; ds.
OS	Homo sapiens.
PN	WO200200928-A2.
PD	03-JAN-2002.
PF	02-JUL-2001; 2001WO-EP07537.
PR	30-JUN-2000; 2000DE-1032529.
PR	01-SEP-2000; 2000DE-1043826.
PA	(EPIG-) EPIGENOMICS AG.
P1	Olek A, Piepenbrock C, Berlin K;
P1	WP1: 2002-130909/17.
P7	Nucleic acid comprising fragment of chemically modified gene, useful
P7	for diagnosis and treatment of diseases associated with abnormal
XX	cytosine methylation -
PS	Claim 1; SEQ ID NO 1170; 32pp + Sequence Listing; German.
XX	
XX	The present invention provides a number of human immune system associated
CC	genes which are modified by the methylation of cytosines. The sequences
CC	can be used in the diagnosis and treatment of immune system disorders,
CC	including eye diseases such as retinopathy, neovascular glaucoma and
CC	macular degeneration, arteriosclerosis, anaemia, cancer, acute myeloid
CC	leukemia, Alzheimer's disease, AIDS, epilepsy, neurofibromatosis,
CC	

CC rheumatoid arthritis, psoriasis and inflammatory/ulcerative bowel
CC diseases. The present sequence is a gene of the invention.

XX Sequence 6183 BP; 1584 A; 230 C; 1532 G; 2837 T; 0 other;

Query Match 3.7%; Score 60.6; DB 24; Length 6183;

Best Local Similarity 88.0%; Pred. No. 9.6e-07;

Matches 66; Conservative 0; Mismatches 9; Indels 0; Gaps 0;

Qy 205 AATAGCTTTACAGTACATTCAAAACACACACACACACACACACACACACAC 264
Db 3373 AAAACCTTACTACTACATAATAAAAACACACACACACACACACACACACAC 3314
Oy 265 ACACACACACACACA 279
Db 3313 ACACACACACACACA 3299

RESULT 8

AAK70103 standard; DNA: 23885 BP.

XX AAK70103;

XX 06-NOV-2001 (first entry)

DE Human immune/haematopoietic antigen genomic sequence SEQ ID NO:24915.

XX Human; immune; haematopoietic; immune/haematopoietic antigen; cancer;

XX cytosolic; gene therapy; vaccine; metastasis; ds.

XX Homo sapiens.

XX WO200157182-A2.

XX 09-AUG-2001.

XX 17-JAN-2001; 2001WO-US01354.

XX
PR 31-JAN-2000; 2000US-0179065.
PR 04-FEB-2000; 2000US-0180628.
PR 24-FEB-2000; 2000US-0184664.
PR 02-MAR-2000; 2000US-0186350.
PR 16-MAR-2000; 2000US-0189874.
PR 17-MAR-2000; 2000US-0190076.
PR 18-APR-2000; 2000US-0198123.
PR 19-MAY-2000; 2000US-0205515.
PR 07-JUN-2000; 2000US-0209467.
PR 28-JUN-2000; 2000US-0214886.
PR 30-JUN-2000; 2000US-0215135.
PR 07-JUL-2000; 2000US-0216647.
PR 07-JUL-2000; 2000US-0216880.
PR 11-JUL-2000; 2000US-0217487.
PR 11-JUL-2000; 2000US-0217496.
PR 14-JUL-2000; 2000US-0218290.
PR 26-JUL-2000; 2000US-0220963.
PR 26-JUL-2000; 2000US-0220964.
PR 14-AUG-2000; 2000US-0224518.
PR 14-AUG-2000; 2000US-0224519.
PR 14-AUG-2000; 2000US-0225213.
PR 14-AUG-2000; 2000US-0225214.
PR 14-AUG-2000; 2000US-0225266.
PR 14-AUG-2000; 2000US-0225267.
PR 14-AUG-2000; 2000US-0225268.
PR 14-AUG-2000; 2000US-0225270.
PR 14-AUG-2000; 2000US-0225447.
PR 14-AUG-2000; 2000US-0225757.
PR 14-AUG-2000; 2000US-0225758.
PR 14-AUG-2000; 2000US-0225759.
PR 18-AUG-2000; 2000US-0226279.
PR 22-AUG-2000; 2000US-0226681.
PR 22-AUG-2000; 2000US-0226866.
PR 22-AUG-2000; 2000US-0227182.

PR 23-AUG-2000; 2000US-0227009.
PR 30-AUG-2000; 2000US-0228924.
PR 01-SEP-2000; 2000US-0229287.
PR 01-SEP-2000; 2000US-0229343.
PR 01-SEP-2000; 2000US-0229344.
PR 01-SEP-2000; 2000US-0229345.
PR 05-SEP-2000; 2000US-0229509.
PR 05-SEP-2000; 2000US-0229513.
PR 06-SEP-2000; 2000US-0230437.
PR 08-SEP-2000; 2000US-0230438.
PR 08-SEP-2000; 2000US-0231242.
PR 08-SEP-2000; 2000US-0231243.
PR 08-SEP-2000; 2000US-0231244.
PR 08-SEP-2000; 2000US-0231413.
PR 08-SEP-2000; 2000US-0231414.
PR 08-SEP-2000; 2000US-0232080.
PR 08-SEP-2000; 2000US-0232081.
PR 12-SEP-2000; 2000US-0231968.
PR 14-SEP-2000; 2000US-0232397.
PR 14-SEP-2000; 2000US-0232398.
PR 14-SEP-2000; 2000US-0232399.
PR 14-SEP-2000; 2000US-0232400.
PR 14-SEP-2000; 2000US-0232401.
PR 14-SEP-2000; 2000US-0233063.
PR 14-SEP-2000; 2000US-0233064.
PR 14-SEP-2000; 2000US-0233065.
PR 21-SEP-2000; 2000US-0234223.
PR 21-SEP-2000; 2000US-0234224.
PR 25-SEP-2000; 2000US-0234597.
PR 25-SEP-2000; 2000US-0234998.
PR 26-SEP-2000; 2000US-0235834.
PR 27-SEP-2000; 2000US-0235834.
PR 27-SEP-2000; 2000US-0235836.
PR 29-SEP-2000; 2000US-0236327.
PR 29-SEP-2000; 2000US-0236367.
PR 29-SEP-2000; 2000US-0236368.
PR 29-SEP-2000; 2000US-0236369.
PR 29-SEP-2000; 2000US-0236370.
PR 02-OCT-2000; 2000US-0236802.
PR 02-OCT-2000; 2000US-0237037.
PR 02-OCT-2000; 2000US-0237038.
PR 02-OCT-2000; 2000US-0237039.
PR 02-OCT-2000; 2000US-0237040.
PR 13-OCT-2000; 2000US-0239935.
PR 13-OCT-2000; 2000US-0239937.
PR 20-OCT-2000; 2000US-0240960.
PR 20-OCT-2000; 2000US-0241221.
PR 20-OCT-2000; 2000US-0241785.
PR 20-OCT-2000; 2000US-0241786.
PR 20-OCT-2000; 2000US-0241787.
PR 20-OCT-2000; 2000US-0241808.
PR 20-OCT-2000; 2000US-0241809.
PR 20-OCT-2000; 2000US-0241826.
PR 01-NOV-2000; 2000US-0244617.
PR 08-NOV-2000; 2000US-0246474.
PR 08-NOV-2000; 2000US-0246475.
PR 08-NOV-2000; 2000US-0246476.
PR 08-NOV-2000; 2000US-0246477.
PR 08-NOV-2000; 2000US-0246478.
PR 08-NOV-2000; 2000US-0246523.
PR 08-NOV-2000; 2000US-0246524.
PR 08-NOV-2000; 2000US-0246525.
PR 08-NOV-2000; 2000US-0246526.
PR 08-NOV-2000; 2000US-0246527.
PR 08-NOV-2000; 2000US-0246528.
PR 08-NOV-2000; 2000US-0246532.
PR 08-NOV-2000; 2000US-0246610.
PR 08-NOV-2000; 2000US-0246619.
PR 08-NOV-2000; 2000US-0246613.
PR 08-NOV-2000; 2000US-0246613.
PR 17-NOV-2000; 2000US-0249207.
PR 17-NOV-2000; 2000US-0249208.
PR 17-NOV-2000; 2000US-0249209.

XX Homo sapiens.
OS WO200164835-A2.
PN 07-SEP-2001.
XX 26-FEB-2001; 2001WO-US04927.
XX 28-FEB-2000; 2000US-0515126.
PR 18-MAY-2000; 2000US-0577409.
XX (HYSE-) HYSEQ INC.
XX Tang YF, Liu C, Drmanac RT;
PI WPI: 2001-514838/56.
DR P-PSDB: AA010445.
XX
XX Isolated nucleic acids and polypeptides, useful for preventing
PT diagnosing and treating e.g. leukaemia, inflammation and immune
PT disorders -
XX
XX Claim 1: SEQ ID NO 10436; 1399pp + Sequence Listing; English.
XX
XX The invention relates to human polynucleotides (AA179941-AA193841) and
CC the encoded proteins (AA000010-AA013910) that exhibit activity elating to
CC cytokine, cell proliferation or cell differentiation or which may induce
CC production of other cytokines in other cell populations. The
CC polynucleotides and polypeptides are useful in gene therapy, vaccines or
CC peptide therapy. The polypeptides have various cytokine-like activities,
CC e.g. stem cell growth factor activity, haematopoiesis regulating
CC activity, tissue growth factor activity, immunomodulatory activity and
CC activin/inhibin activity and may be useful in the diagnosis and/or
CC treatment of cancer, leukaemia, nervous system disorders, arthritis and
CC inflammation.
CC Note: The sequence data for this patent did not form part of the printed
CC specification, but was obtained in electronic format directly from WIPO
CC at ftp.wipo.int/pub/published_pcl_sequences.
XX
XX Sequence 415 BP; 130 A; 109 C; 63 G; 112 T; 1 other;
S0
Query Match 3.6%; Score 60.4; DB 22; Length 415;
Best Local Similarity 74.5%; Pred. No. 2.4e-07;
Matches 76; Conservative 0; Mismatches 26; Indels 0; Gaps 0;
QY 187 CTCACGTATGAGACATAGCTTACAGCTACATTTCAAAACACACACACACAC 246
DB 183 CACCTGTCTACTCATATATAGCTGAAACCTCTTAAGCACACACACACACACAC 242
QY 247 ACACACACACACACACACACACACACACACATGTTTCTT 288
DB 243 ACACACACACACACACACACACACACACACACCTTATT 284
RESULT 11
AAK81626
ID AAK81626 standard; DNA; 8095 BP.
XX
XX AAK81626;
XX
XX 07-NOV-2001 (first entry)
XX
XX Human immune/haematopoietic antigen genomic sequence SEQ ID NO:36438.
XX
XX Human: immune: haematopoietic; immune/haematopoietic antigen; cancer;
KW cytostatic; gene therapy; vaccine; metastasis; ds.
XX
XX Homo sapiens.
XX
XX WO200157182-A2.
XX
XX 09-AUG-2001.

XX
PF 17-JAN-2001; 2001WO-US01354.
XX
XX 31-JAN-2000; 2000US-0179065.
PR 04-FEB-2000; 2000US-0180628.
PR 24-FEB-2000; 2000US-0184664.
PR 02-MAR-2000; 2000US-0186350.
PR 16-MAR-2000; 2000US-0189874.
PR 17-MAR-2000; 2000US-0190076.
PR 18-APR-2000; 2000US-0198123.
PR 19-MAY-2000; 2000US-0205151.
PR 07-JUN-2000; 2000US-0209467.
PR 28-JUN-2000; 2000US-0214886.
PR 30-JUN-2000; 2000US-0215135.
PR 07-JUL-2000; 2000US-0216647.
PR 07-JUL-2000; 2000US-0216880.
PR 11-JUL-2000; 2000US-0217487.
PR 14-JUL-2000; 2000US-0218290.
PR 26-JUL-2000; 2000US-0220963.
PR 26-JUL-2000; 2000US-0220964.
PR 14-AUG-2000; 2000US-0224518.
PR 14-AUG-2000; 2000US-0224519.
PR 14-AUG-2000; 2000US-0225213.
PR 14-AUG-2000; 2000US-0225214.
PR 14-AUG-2000; 2000US-0225266.
PR 14-AUG-2000; 2000US-0225267.
PR 14-AUG-2000; 2000US-0225268.
PR 14-AUG-2000; 2000US-0225270.
PR 14-AUG-2000; 2000US-0225447.
PR 14-AUG-2000; 2000US-0225757.
PR 14-AUG-2000; 2000US-0225758.
PR 14-AUG-2000; 2000US-0225759.
PR 18-AUG-2000; 2000US-0226279.
PR 22-AUG-2000; 2000US-0226811.
PR 22-AUG-2000; 2000US-0226868.
PR 22-AUG-2000; 2000US-0227182.
PR 23-AUG-2000; 2000US-0227009.
PR 30-AUG-2000; 2000US-0228924.
PR 01-SEP-2000; 2000US-0228287.
PR 01-SEP-2000; 2000US-0228343.
PR 01-SEP-2000; 2000US-0229344.
PR 01-SEP-2000; 2000US-0229345.
PR 05-SEP-2000; 2000US-0229509.
PR 05-SEP-2000; 2000US-0229513.
PR 06-SEP-2000; 2000US-0230437.
PR 06-SEP-2000; 2000US-0230438.
PR 08-SEP-2000; 2000US-0231242.
PR 08-SEP-2000; 2000US-0231243.
PR 08-SEP-2000; 2000US-0231244.
PR 08-SEP-2000; 2000US-0231413.
PR 08-SEP-2000; 2000US-0231414.
PR 08-SEP-2000; 2000US-0232080.
PR 08-SEP-2000; 2000US-0232081.
PR 12-SEP-2000; 2000US-0231968.
PR 14-SEP-2000; 2000US-0232397.
PR 14-SEP-2000; 2000US-0232398.
PR 14-SEP-2000; 2000US-0232399.
PR 14-SEP-2000; 2000US-0232400.
PR 14-SEP-2000; 2000US-0232401.
PR 14-SEP-2000; 2000US-0233063.
PR 14-SEP-2000; 2000US-0233064.
PR 21-SEP-2000; 2000US-0233065.
PR 21-SEP-2000; 2000US-0234223.
PR 21-SEP-2000; 2000US-0234274.
PR 25-SEP-2000; 2000US-0234997.
PR 25-SEP-2000; 2000US-0234998.
PR 26-SEP-2000; 2000US-0235484.
PR 27-SEP-2000; 2000US-0235834.
PR 27-SEP-2000; 2000US-0235836.
PR 29-SEP-2000; 2000US-0236327.
PR 29-SEP-2000; 2000US-0236367.
PR 29-SEP-2000; 2000US-0236368.

PR	29-SEP-2000	2000US-0236369.
PR	29-SEP-2000	2000US-0236370.
PR	02-OCT-2000	2000US-0236802.
PR	02-OCT-2000	2000US-0237037.
PR	02-OCT-2000	2000US-0237038.
PR	02-OCT-2000	2000US-0237039.
PR	02-OCT-2000	2000US-0237040.
PR	13-OCT-2000	2000US-0239935.
PR	13-OCT-2000	2000US-0239937.
PR	20-OCT-2000	2000US-0240960.
PR	20-OCT-2000	2000US-0241221.
PR	20-OCT-2000	2000US-0241785.
PR	20-OCT-2000	2000US-0241786.
PR	20-OCT-2000	2000US-0241787.
PR	20-OCT-2000	2000US-0241808.
PR	20-OCT-2000	2000US-0241809.
PR	20-OCT-2000	2000US-0241826.
PR	01-NOV-2000	2000US-0244617.
PR	08-NOV-2000	2000US-0246474.
PR	08-NOV-2000	2000US-0246475.
PR	08-NOV-2000	2000US-0246476.
PR	08-NOV-2000	2000US-0246477.
PR	08-NOV-2000	2000US-0246478.
PR	08-NOV-2000	2000US-0246523.
PR	08-NOV-2000	2000US-0246524.
PR	08-NOV-2000	2000US-0246525.
PR	08-NOV-2000	2000US-0246526.
PR	08-NOV-2000	2000US-0246527.
PR	08-NOV-2000	2000US-0246528.
PR	08-NOV-2000	2000US-0246532.
PR	08-NOV-2000	2000US-0246539.
PR	08-NOV-2000	2000US-0246610.
PR	08-NOV-2000	2000US-0246611.
PR	08-NOV-2000	2000US-0246613.
PR	17-NOV-2000	2000US-0249207.
PR	17-NOV-2000	2000US-0249208.
PR	17-NOV-2000	2000US-0249209.
PR	17-NOV-2000	2000US-0249210.
PR	17-NOV-2000	2000US-0249211.
PR	17-NOV-2000	2000US-0249212.
PR	17-NOV-2000	2000US-0249213.
PR	17-NOV-2000	2000US-0249214.
PR	17-NOV-2000	2000US-0249215.
PR	17-NOV-2000	2000US-0249216.
PR	17-NOV-2000	2000US-0249217.
PR	17-NOV-2000	2000US-0249218.
PR	17-NOV-2000	2000US-0249244.
PR	17-NOV-2000	2000US-0249245.
PR	17-NOV-2000	2000US-0249264.
PR	17-NOV-2000	2000US-0249265.
PR	17-NOV-2000	2000US-0249267.
PR	17-NOV-2000	2000US-0249299.
PR	17-NOV-2000	2000US-0249300.
PR	01-DEC-2000	2000US-0250160.
PR	01-DEC-2000	2000US-0250391.
PR	05-DEC-2000	2000US-0251030.
PR	05-DEC-2000	2000US-0251988.
PR	05-DEC-2000	2000US-0256719.
PR	06-DEC-2000	2000US-0251479.
PR	08-DEC-2000	2000US-0251856.
PR	08-DEC-2000	2000US-0251868.
PR	08-DEC-2000	2000US-0251869.
PR	08-DEC-2000	2000US-0251989.
PR	11-DEC-2000	2000US-0251990.
PR	05-JAN-2001	2001US-0259678.
XX		
PA	(HUMA-)	HUMAN GENOME SCI INC.
XX		
PI	Rosen CA,	Barash SC, Ruben SM;
XX		
DR	WPI,	2001-483426/52.
XX		

PT	Nucleic acids encoding human immune/hematopoietic antigen polypeptides, useful for preventing, diagnosing and/or treating cancers and metastasis -
PT	
XX	Disclosure; SEQ ID NO 36438; 3071pp + Sequence Listing; English.
PS	
XX	
CC	AAK54951 to AAK64702 encode the human immune/hematopoietic antigen (I)
CC	amino acid sequences given in AAM82170 to AAM91921. (I) have cytostatic
CC	activity, and can be used in gene therapy and vaccine production. (I)
CC	proteins and polynucleotides may be used in the prevention, diagnosis and
CC	treatment of diseases associated with inappropriate (I) expression. For
CC	example, they may be used to treat disorders associated with decreased
CC	expression by rectifying mutations or deletions in a patient's genome
CC	that affect the activity of (I) by expressing inactive proteins or to
CC	supplement the patients own production of (I). Additionally, (I)
CC	polynucleotides may be used to produce the secreted (I) by inserting the
CC	nucleic acids into a host cell and culturing the cell to express the
CC	protein. (I) proteins and polynucleotides may be used to prevent,
CC	diagnose and treat immune/hematopoietic-related diseases, especially
CC	cancers and cancer metastases of hematopoietic-derived cells. AAK64703
CC	to AAK87654 represent human immune/hematopoietic antigen genomic
CC	sequences from the present invention. AAK5942 to AAK54950 and AAM82169
CC	represent sequences used in the exemplification of the present invention.
CC	
XX	
SO	Sequence 8095 BP; 2662 A; 1135 C; 1557 G; 2741 T; 0 other;
	Query Match 3.6%; Score 60.4; DB 22; Length 8095;
	Best Local Similarity 81.4%; Pred. No. 1.3e-05;
	Matches 70; Conservative 0; Mismatches 16; Indels 0; Gaps 0;
OY	194 GATATGGACAGAAATGCTTACACTCATTTCAAACACACACACACACACACACA 253
Db	7247 CAAGAAGAGGGATGTGAGATTCCCCCTTACATACACACACACACACACACACACACA 7306
OY	254 CACACACACACACACACACACACACA 279
DB	7307 CACACACACACACACACACACACACA 7332
RESULT 12	
ID	AAC91904 standard; cDNA; 4858 BP.
XX	
AC	AAC91904;
XX	
DT	19-MAR-2001 (first entry)
XX	
DE	Murine A259 coding sequence.
XX	
KM	Murine; A259: integrin alpha subunit; fibrosis; liver disease; lung;
KW	kidney; cancer; osteoporosis; acute myeloid leukaemia; HIV infection;
XX	rheumatoid arthritis; ss.
OS	
XX	Mus sp.
FT	
PH	Key Location/Qualifiers
FT	CDS 28..3594
FT	/tag= a
FT	/product= "Murine A259"
FT	sig_peptide 28..93
FT	/tag= b
FT	mat_peptide 94..3591
FT	/tag= c
PN	
XX	WO200073339-A1.
PD	
XX	07-DEC-2000.
PF	15-MAY-2000; 2000WO-US13262.
XX	
PR	28-MAY-1999; 99US-0322790.
PR	27-APR-2000; 2000US-0561263.
XX	

PA (MILL-) MILLENNIUM PHARM INC.
XX
XX Pan Y, Lora JM;
XX
XX WPI: 2001-041142/05.
DR P-PSDB: AAB50087.
XX
XX Nucleic acid encoding alpha-integrin subunits, useful for treatment and
PT diagnosis of fibrosis, e.g. of the liver
XX
XX Claim 1; Fig 5; 164pp; English.
XX
XX The present sequence is the coding sequence for murine integrin alpha
CC subunit, A259. A259 is homologous with the alpha and alpha10 integrin
CC subunits and is overexpressed in fibrosis. A259 is implicated in
CC regulation of proliferation, differentiation and/or function of many
CC different cell types. Inhibitors of A259 activity are useful for the
CC treatment of liver disease, particularly fibrosis, and also fibrosis in
CC other organs (specifically lung and kidney). In addition, A259 can be
CC used for treatment and prevention of cancer, osteoporosis, acute myeloid
CC leukaemia, HIV infection, and rheumatoid arthritis.
XX
XX Sequence 4858 BP; 1184 A; 1397 C; 1296 G; 981 T; 0 other;
SQ
Query Match 3.6%; Score 60.2; DB 22; Length 4858;
Best Local Similarity 68.6%; Pred. No. 1.1e-06;
Matches 83; Conservative 0; Mismatches 38; Indels 0; Gaps 0;
OY 179 CCCGGAGCTCCACTGATATGGACAGATAGCTTTACAGCTACATTCAAAACACACACAC 238
DB 4398 CCGGTAGATGCACTGTGATGGAGCCCTGTCTTACATCAGCAGCTACACACACAC 4457
OY 239 ACACACACACACACACACACACACACACACACATGTTTCTCCCTCCCTCC 298
DB 4458 ACATC 4517
OY 299 A 299
DB 4518 A 4518
RESULT 13
AAS16874
ID AAS16874 standard; cDNA; 4858 BP.
XX
AC AAS16874;
XX
DT 14-FEB-2002 (first entry)
XX
DE Murine A259 cDNA.
XX
KW Mouse; A259, integrin alpha subunit; integrin alpha 10; secreted protein;
KW liver disease; fibrosis; lung; kidney; bone associated disorder; blood;
KW cartilage associated disorder; haematopoietic disorder; bone marrow; ss;
KW immune related disease; apoptotic disorder; neuronal tissue disease;
KW neurodegenerative disease; gene therapy; cancer; cytostatic; osteopathic;
KW nephrotropic; immunomodulator; anti-inflammatory; neuroprotective;
KW antiarthritic; antinaemic; antiallergic; antialastmatic; dermatological;
KW antidiabetic; anticonvulsant; antiparkinsonian.
XX
XX Mus musculus.
XX
OS
XX
FH Key location/qualifiers
FT CDS 28..3594
FT
FT /*tag= a
FT /product= "Mouse A259"
FT /note= "This coding sequence is specifically claimed"
FT sig_peptide 28..93
FT /*tag= b
FT mat_peptide 94..3591
FT /*tag= c
FT /product= "Mature murine A259"
XX

PN WO200181414-A2.
XX
XX 01-NOV-2001.
XX
XX 27-APR-2001; 2001MO-US13516.
XX
XX 27-APR-2000; 2000US-0561263.
XX
XX (MILL-) MILLENNIUM PHARM INC.
PA
XX Pan Y, Lora J;
XX
XX WPI: 2002-041397/05.
DR P-PSDB: AAU10552.
XX
XX New A259 nucleic acids and polypeptides, which comprise integrin alpha
PT subunit, useful for diagnosing, preventing or treating e.g. liver
PT disease, kidney or lung fibrosis, cancers, blood disorders or immune
PT related diseases
XX
XX Claim 2; Fig 5; 168pp; English.
XX
XX The invention relates to human and murine A259 nucleic acid molecules
CC which encode secreted proteins with homology to integrin alpha subunits,
CC specifically to integrin alpha 10. The A259 polypeptide and nucleic acid
CC are useful for treating liver disease or fibrosis, particularly kidney
CC fibrosis or lung fibrosis. The A259 polypeptide and nucleic acid are also
CC useful for diagnosing, preventing or treating cartilage and bone
CC associated disorders (such as bone cancer, achondroplasia, myeloma,
CC fibrous dysplasia, scoliosis, osteoarthritis, osteosarcoma and
CC osteoporosis), bone marrow, blood and haematopoietic disorders (such as
CC acute myeloid leukaemia, haemophilia, anaemia and thalassemia), immune
CC related diseases (such as HIV, viral infections, cancers, T cell
CC autoimmune disorders e.g. AIDS, and allergic inflammatory disorders e.g.
CC asthma and psoriasis), apoptotic disorders (such as systemic lupus
CC erythematosus and insulin dependent diabetes mellitus), diseases of the
CC neuronal tissues (such as epilepsy and muscular dystrophy) and
CC neurodegenerative diseases (such as Parkinson's disease and Huntington's
CC disease). This sequence represents cDNA encoding the murine A259
CC polypeptide.
XX
XX Sequence 4858 BP; 1184 A; 1397 C; 1296 G; 981 T; 0 other;
SQ
Query Match 3.6%; Score 60.2; DB 24; Length 4858;
Best Local Similarity 68.6%; Pred. No. 1.1e-06;
Matches 83; Conservative 0; Mismatches 38; Indels 0; Gaps 0;
OY 179 CCCGGAGCTCCACTGATATGGACAGATAGCTTTACAGCTACATTCAAAACACACACAC 238
DB 4398 CCGGTAGATGCACTGTGATGGAGCCCTGTCTTACATCAGCAGCTACACACACAC 4457
OY 239 ACACACACACACACACACACACACACACACACATGTTTCTCCCTCCCTCC 298
DB 4458 ACATC 4517
OY 299 A 299
DB 4518 A 4518
RESULT 14
AAA39622
ID AAA39622 standard; DNA; 1787 BP.
XX
AC AAA39622;
XX
DT 01-SEP-2000 (first entry)
XX
DE Murine stomatin DNA.
XX
KW Stomatin; murine; treatment; pain; analgesic; anesthetic; allodynia;
KW diagnostic; ss.
XX

29-SEP-2000; 2000US-0236368.
 PR 29-SEP-2000; 2000US-0236369.
 PR 29-SEP-2000; 2000US-0236370.
 PR 02-OCT-2000; 2000US-0236802.
 PR 02-OCT-2000; 2000US-0237037.
 PR 02-OCT-2000; 2000US-0237038.
 PR 02-OCT-2000; 2000US-0237039.
 PR 02-OCT-2000; 2000US-0237040.
 PR 13-OCT-2000; 2000US-0239935.
 PR 13-OCT-2000; 2000US-0239937.
 PR 20-OCT-2000; 2000US-0240960.
 PR 20-OCT-2000; 2000US-0241221.
 PR 20-OCT-2000; 2000US-0241785.
 PR 20-OCT-2000; 2000US-0241786.
 PR 20-OCT-2000; 2000US-0241787.
 PR 20-OCT-2000; 2000US-0241808.
 PR 20-OCT-2000; 2000US-0241809.
 PR 20-OCT-2000; 2000US-0241826.
 PR 01-NOV-2000; 2000US-0244617.
 PR 08-NOV-2000; 2000US-0244617.
 PR 08-NOV-2000; 2000US-0246475.
 PR 08-NOV-2000; 2000US-0246475.
 PR 08-NOV-2000; 2000US-0246476.
 PR 08-NOV-2000; 2000US-0246477.
 PR 08-NOV-2000; 2000US-0246478.
 PR 08-NOV-2000; 2000US-0246523.
 PR 08-NOV-2000; 2000US-0246524.
 PR 08-NOV-2000; 2000US-0246525.
 PR 08-NOV-2000; 2000US-0246526.
 PR 08-NOV-2000; 2000US-0246527.
 PR 08-NOV-2000; 2000US-0246528.
 PR 08-NOV-2000; 2000US-0246532.
 PR 08-NOV-2000; 2000US-0246609.
 PR 08-NOV-2000; 2000US-0246610.
 PR 08-NOV-2000; 2000US-0246611.
 PR 17-NOV-2000; 2000US-0249207.
 PR 17-NOV-2000; 2000US-0249208.
 PR 17-NOV-2000; 2000US-0249209.
 PR 17-NOV-2000; 2000US-0249210.
 PR 17-NOV-2000; 2000US-0249211.
 PR 17-NOV-2000; 2000US-0249212.
 PR 17-NOV-2000; 2000US-0249213.
 PR 17-NOV-2000; 2000US-0249214.
 PR 17-NOV-2000; 2000US-0249215.
 PR 17-NOV-2000; 2000US-0249216.
 PR 17-NOV-2000; 2000US-0249217.
 PR 17-NOV-2000; 2000US-0249218.
 PR 17-NOV-2000; 2000US-0249219.
 PR 17-NOV-2000; 2000US-0249224.
 PR 17-NOV-2000; 2000US-0249245.
 PR 17-NOV-2000; 2000US-0249264.
 PR 17-NOV-2000; 2000US-0249265.
 PR 17-NOV-2000; 2000US-0249297.
 PR 17-NOV-2000; 2000US-0249299.
 PR 01-DEC-2000; 2000US-0249300.
 PR 01-DEC-2000; 2000US-0250160.
 PR 01-DEC-2000; 2000US-0250391.
 PR 05-DEC-2000; 2000US-0251030.
 PR 05-DEC-2000; 2000US-0251988.
 PR 05-DEC-2000; 2000US-0251988.
 PR 05-DEC-2000; 2000US-0251988.
 PR 08-DEC-2000; 2000US-0251856.
 PR 08-DEC-2000; 2000US-0251856.
 PR 08-DEC-2000; 2000US-0251869.
 PR 08-DEC-2000; 2000US-0251869.
 PR 08-DEC-2000; 2000US-0251989.
 PR 11-DEC-2000; 2000US-0254097.
 PR 05-JAN-2001; 2001US-0259678.

(HUMA-) HUMAN GENOME SCI INC.

PA Rosen CA, Barash SC, Ruben SM;
 PI
 XX
 DR WPI; 2001-483426/52.

XX Nucleic acids encoding human immune/hematopoietic antigen polypeptides,
 PT useful for preventing, diagnosing and/or treating cancers and
 PT metastasis -
 XX
 PS Disclosure: SEQ ID NO 21626; 3071pp + Sequence Listing; English.
 XX
 CC AAK54951 to AAK64702 encode the human immune/hematopoietic antigen (I)
 CC amino acid sequences given in AAM82170 to AAM91921. (I) have cytostatic
 CC activity, and can be used in gene therapy and vaccine production. (I)
 CC proteins and polynucleotides may be used in the prevention, diagnosis and
 CC treatment of diseases associated with inappropriate (I) expression. For
 CC example, they may be used to treat disorders associated with decreased
 CC expression by rectifying mutations or deletions in a patient's genome
 CC that affect the activity of (I) by expressing inactive proteins or to
 CC supplement the patient's own production of (I). Additionally, (I)
 CC polynucleotides may be used to produce the secreted (I), by inserting
 CC the nucleic acids into a host cell and culturing the cell to express the
 CC protein. (I) proteins and polynucleotides may be used to prevent,
 CC diagnose and treat immune/hematopoietic-related diseases, especially
 CC cancers and cancer metastases of haematopoietic-derived cells. AAK64703
 CC to AAK87694 represent human immune/hematopoietic antigen genomic
 CC sequences from the present invention. AAK54942 to AAK54950 and AAM82169
 CC represent sequences used in the exemplification of the present invention.
 XX
 SQ Sequence 5176 BP; 1472 A; 1210 C; 1170 G; 1324 T; 0 other:
 Query Match 3.6%; Score 59.8; DB 22; Length 5176;
 Best Local Similarity 84.8%; Pred. NO.1.5e-06;
 Matches 67; Conservative 0; Mismatches 12; Indels 0; Gaps 0;
 QY 227 AACACACACACACACACACACACACACACACACACACATGTTTC 286
 DB 819 AACATGTTT 878
 QY 287 TTCCCTCCCTCCACTTCT 305
 DB 879 ACCCTCTCATCCCTTCTCT 897

Search completed: November 6, 2002, 14:16:40
 Job time : 510.209 secs

GenCore version 5.1.3
Copyright (c) 1993 - 2002 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: November 6, 2002, 14:20:41 ; Search time 2257.42 seconds
(without alignments)
15388.369 Million cell updates/sec

Title: US-09-833-740-1_COPY_1_1660

Perfect score: 1660
Sequence: 1 aggttaccgcatttgaca.....gaagagactgtgcaaca 1660

Scoring table: IDENTITY-NUC
Gapop 10.0, Gapext 1.0

Searched: 1797656 seqs, 10463268293 residues

Total number of hits satisfying chosen parameters: 3595312

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 08
Maximum Match 100%
Listing first 45 summaries

Database : GenEmbl:*
1: gb_da:*
2: gb_htg:*
3: gb_in:*
4: gb_om:*
5: gb_ov:*
6: gb_pat:*
7: gb_pl:*
8: gb_pr:*
9: gb_ro:*
10: gb_sts:*
11: gb_sy:*
12: gb_un:*
13: gb_vl:*
14: gb_vl:*
15: em_fun:*
16: em_fun:*
17: em_hum:*
18: em_hum:*
19: em_mu:*
20: em_mu:*
21: em_or:*
22: em_ov:*
23: em_pat:*
24: em_ph:*
25: em_pl:*
26: em_ro:*
27: em_sts:*
28: em_un:*
29: em_vl:*
30: em_vl:*
31: em_htg_hum:*
32: em_htg_inv:*
33: em_htg_inv:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Query	Match Length	ID	Description
------------	-------	--------------	----	-------------

1	1658.8	99.9	2170	6	AX354807	AX354807 Sequence
2	1658.8	99.9	3130	10	AF338224	AF338224 Mus muscu
3	1601.2	96.5	207160	2	AC016464	AC016464 Mus muscu
c	1601.2	96.5	211173	2	AL646097	AL646097 Mus muscu
5	175	10.5	493	6	AX354812	AX354812 Sequence
c	101.6	6.1	198271	2	AC026591	AC026591 Homo sapi
c	101.6	6.1	210719	2	AC087646	AC087646 Homo sapi
8	101	6.1	213359	2	AC069006	AC069006 Homo sapi
9	77.4	4.7	49665	2	AC100859	AC100859 Homo sapi
10	73.2	4.4	102783	2	AC096120	AC096120 Rattus no
c	72.2	4.3	171322	2	AC095179	AC095179 Rattus no
c	72.2	4.3	72930	2	AC101368	AC101368 Mus muscu
13	71.2	4.3	17529	10	AF417202	AF417202 Mus muscu
14	68.8	4.1	176047	2	AC099092	AC099092 Rattus no
c	68.2	4.1	184550	2	AC095653	AC095653 Rattus no
15	68	4.1	177026	2	AC097947	AC097947 Rattus no
16	67.6	4.1	269035	2	AC093941	AC093941 Rattus no
c	67.4	4.1	74978	2	AC101567	AC101567 Mus muscu
18	67.4	4.0	162540	2	AC087898	AC087898 Mus muscu
19	67.2	4.0	179604	2	AC106235	AC106235 Rattus no
20	67.2	4.0	162540	2	AC099716	AC099716 Mus muscu
21	67.2	4.0	179604	2	AC093949	AC093949 Rattus no
22	67	4.0	155258	2	AC023605	AC023605 Mus muscu
c	67	4.0	292943	2	AC023605	AC023605 Mus muscu
23	66.8	4.0	447	11	G40821	G40821 Zebrafish
24	66.8	4.0	183100	2	AC097616	AC097616 Rattus no
25	66.8	4.0	184884	2	AC105552	AC105552 Rattus no
c	66.6	4.0	284	8	PAB292695	PAB292695 Picea abi
27	66.6	4.0	646	11	G40571	G40571 Zebrafish
28	66.6	4.0	16430	10	AB06058852	AB06058852 Mus muscu
c	66.6	4.0	121868	2	AC103157	AC103157 Rattus no
30	66.6	4.0	156424	2	AC095092	AC095092 Rattus no
31	66.6	4.0	152316	10	AC007844	AC007844 Mus muscu
32	66.4	4.0	173500	2	AC097832	AC097832 Rattus no
c	66.4	4.0	183665	2	AL645600	AL645600 Mus muscu
34	66.4	4.0	216819	2	AL645725	AL645725 Mus muscu
c	66.4	4.0	140115	2	AC091470	AC091470 Mus muscu
35	66.2	4.0	163177	2	AC106056	AC106056 Rattus no
c	66.2	4.0	168833	2	AC021706	AC021706 Homo sapi
37	66.2	4.0	198578	2	AC087138	AC087138 Mus muscu
38	66.2	4.0	221511	2	AC087556	AC087556 Mus muscu
40	66.2	4.0	222817	2	AL662803	AL662803 Mus muscu
41	66.2	4.0	364	11	HS210WES	HS210WES H. sapiens
42	66	4.0	107110	2	AC103028	AC103028 Rattus no
c	65.8	4.0	320	11	G45569	G45569 Zebrafish
43	65.8	4.0	158949	2	AC094758	AC094758 Rattus no
c	65.8	4.0	158949	2	AC094758	AC094758 Rattus no

ALIGNMENTS

RESULT 1
AX354807
LOCUS AX354807
DEFINITION Sequence 1 from Patent WO0179290.
ACCESSION AX354807
VERSION AX354807.1 GI:18619538
KEYWORDS
SOURCE
ORGANISM
synthetic construct.
synthetic construct.
artificial sequence.
REFERENCE
1 (sites)
AUTHORS Drucker,D.J. and Lovshin,J.A.
TITLE GIP-2 receptor gene promoter and uses thereof
JOURNAL Patent: WO 0179290-A1 25-Oct-2001;
1149336 ONTARIO INC. (CA)
FEATURES
source
1..2170
/organism="synthetic construct"
/db_xref="taxon:32630"
/note="Recombinant DNA expression construct"

BASE COUNT	574 a	500 c	531 g	540 t	5 others
------------	-------	-------	-------	-------	----------

Query Match	99.9%	Score 1658.8;	DB 6;	Length 2170;
Best Local Similarity	100.0%	Pred. No. 0;		
Matches 1660;	Conservative 0;	Mismatches 0;	Indels 0;	Gaps 0.

Qy	1	AGGTTACCGGATTTTGACATAGTGGCATTCGCGCCACGGGTGCGAGTATGAAGT	60
Db	1	AGGTTTACCGGATTTTGACATAGTGGCATTCGCGCCACGGGTGCGAGTATGAAGT	60
Qy	61	GACCAAGGCAAGTCCTTTACAGGGGGAGAAAAATCAGAAAAAAATTTTAGAAGCAT	120
Db	61	GACCAAGGCAAGTCCTTTACAGGGGGAGAAAAATCAGAAAAAAATTTTAGAAGCAT	120
Qy	121	TTTCAAGAAAGCAAGATGGAATATTGTACAAAAACGTCGTTCTCCCCACATGGACC	180
Db	121	TTTCAAGAAAGCAAGATGGAATATTGTACAAAAACGTCGTTCTCCCCACATGGACC	180
Qy	181	CGGAGAGTCACAGTATATGACAGAAATAGCTTTACAGCTACATTCMAAACACACACAC	240
Db	181	CGGAGAGTCACAGTATATGACAGAAATAGCTTTACAGCTACATTCMAAACACACACAC	240
Qy	241	ACACACACACACACACACACACACACACACACACATGTTTTCTCCCTCCCTCCAC	300
Db	241	ACACACACACACACACACACACACACACACACACATGTTTTCTCCCTCCCTCCAC	300
Qy	301	TTCCGCCCATTCCTCGTGTGTCGCCAAGAGATACCATATTGACTGAGAAATCACACAC	360
Db	301	TTCCGCCCATTCCTCGTGTGTCGCCAAGAGATACCATATTGACTGAGAAATCACACAC	360
Qy	361	CATAAAGCCCCATCTGGGAGCCATTTCCAGACTGATCTTTTATCATTTAAGGTTGAAT	420
Db	361	CATAAAGCCCCATCTGGGAGCCATTTCCAGACTGATCTTTTATCATTTAAGGTTGAAT	420
Qy	421	CTTGCCACGTGGGTTTTTAAGGTTTTTAAGGATTTTTTATCAGCGCACTCACGCTT	480
Db	421	CTTGCCACGTGGGTTTTTAAGGTTTTTAAGGATTTTTTATCAGCGCACTCACGCTT	480
Qy	481	CCCTGTGAATGTTCAAGAAATCTACTGGGCTTGTCACGCTAATGGAAATGATATGGTTTG	540
Db	481	CCCTGTGAATGTTCAAGAAATCTACTGGGCTTGTCACGCTAATGGAAATGATATGGTTTG	540
Qy	541	ACTTAAATGTGAAGAAAAAGAAAAAGAGGGGGAAAAAGAGGGGAGAAAAAGAGGGGAAG	600
Db	541	ACTTAAATGTGAAGAAAAAGAAAAAGAGGGGGAAAAAGAGGGGAGAAAAAGAGGGGAAG	600
Qy	601	GGAAAACTGCCCTTTTATGCTATTGCTACTCTAACAATTTTGTCCTCACCTTCACCTTGG	660
Db	601	GGAAAACTGCCCTTTTATGCTATTGCTACTCTAACAATTTTGTCCTCACCTTCACCTTGG	660
Qy	661	TTCTTCATATGGAAGAACTGGATAGAAAGCTGGAGGACAGCCAGAGGATAGGAGAGTGT	720
Db	661	TTCTTCATATGGAAGAACTGGATAGAAAGCTGGAGGACAGCCAGAGGATAGGAGAGTGT	720
Qy	721	GTCGTGTGTGGGGGGGGTGGGCGCAGAGAGAGACTTAGAGACAGAAAGAGCCCTGTAG	780
Db	721	GTCGTGTGTGGGGGGGGTGGGCGCAGAGAGAGACTTAGAGACAGAAAGAGCCCTGTAG	780
Qy	781	ACAATCATAGCTTTCCTTTAGAGACCCCTAGTGTCTAAGAGAAATGTTCTTAACAGAGTAC	840
Db	781	ACAATCATAGCTTTCCTTTAGAGACCCCTAGTGTCTAAGAGAAATGTTCTAACAGAGTAC	840
Qy	841	TGTGGTCAGCTAGCTGGGCTGGAAAGSCTGGCTTTGCTTTTGTCTTGTCTGTCAGGCTT	900
Db	841	TGTGGTCAGCTAGCTGGGCTGGAAAGSCTGGCTTTGCTTTTGTCTTGTCTGTCAGGCTT	900
Qy	901	GAACAACAACCCCTGTCCTTTGTAACCCCACTATTCTCAGCCCTCAGATGAAGAGTAA	960
Db	901	GAACAACAACCCCTGTCCTTTGTAACCCCACTATTCTCAGCCCTCAGATGAAGAGTAA	960
Qy	961	TGCTACCTTGGAGATACATGATGGGTTCAAGTAACTAGGAGAGGGGTGGAAGGTTTGG	1020
Db	961	TGCTACCTTGGAGATACATGATGGGTTCAAGTAACTAGGAGAGGGGTGGAAGGTTTGG	1020
Qy	1021	TAAACCTAAACTGAACTGGGGTGTGGTTGTAAAGTAGCCATGATCCATTAATAATATC	1080
1.			

[illegible]

mRNA 2617...>2900
/product="glucagon-like peptide-2 receptor"
2835...>2900
CDS /note="GLP-2 receptor"
/codon_start=1
/product="glucagon-like peptide-2 receptor"
/protein_id="AAK63043.1"
/db_xref="GI:14485647"
/translation="MRRLMGPEPPLSLLLVSIKQ"
BASE COUNT 805 a 741 c 769 g 797 t 18 others
ORIGIN

Query Match 99.9%; Score 1658.8; DB 10; Length 3130;
Best Local Similarity 100.0%; Pred. No. 0;
Matches 1660; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 AGCTTATCCGCAATTTTACACTAGATGCAATCCGTCACGCGGTAGCAGTCAAGCT 60
DB 961 AGCTTATCCGCAATTTTACACTAGATGCAATCCGTCACGCGGTAGCAGTCAAGCT 1020
QY 61 GACCAAGGCAAGCTCTTTCAGGGGGAAGAAATCAGAAAAAATTTTGAAGCAT 120
DB 1021 GACCAAGGCAAGCTCTTTCAGGGGGAAGAAATCAGAAAAAATTTTGAAGCAT 1080
QY 121 TTCAAGAACCAAGATGGAATATTTTACAAAACAGGTCTTCTCCCAATGCGACC 180
DB 1081 TTCAAGAACCAAGATGGAATATTTTACAAAACAGGTCTTCTCCCAATGCGACC 1140
QY 181 CGGAGCTCCACTGATATGAGACAGAAATAGCTTTACAGCTACATTCAAAACACACACAC 240
DB 1141 CGGAGCTCCACTGATATGAGACAGAAATAGCTTTACAGCTACATTCAAAACACACACAC 1200
QY 241 AC 300
DB 1201 AC 1260
QY 301 TTCTCCCAATCTCTGTGTGTCACCAAGAGATGACATATTTGACTGTGAATCACAAC 360
DB 1261 TTCTCCCAATCTCTGTGTGTCACCAAGAGATGACATATTTGACTGTGAATCACAAC 1320
QY 361 CATAAAAAGCCCATCTGGAGCCATTTCCAGACATCTTTTATCATTAAGTTTGAATT 420
DB 1321 CATAAAAAGCCCATCTGGAGCCATTTCCAGACATCTTTTATCATTAAGTTTGAATT 1380
QY 421 CTTCGCCAGCTGTGGTTTAAAGTTTAAAGTTTAAAGTTTAAAGTTTAAAGTTTAA 480
DB 1381 CTTCGCCAGCTGTGGTTTAAAGTTTAAAGTTTAAAGTTTAAAGTTTAAAGTTTAA 1440
QY 481 CCTGTGAATGTTCAGAAATCACTGGGCTGTGACGCTAATGGAATGATCTATGTTTG 540
DB 1441 CCTGTGAATGTTCAGAAATCACTGGGCTGTGACGCTAATGGAATGATCTATGTTTG 1500
QY 541 ACTTAAATGTGAAGAAAAAAGAGGGGGAAGAGGAGGAGGAGGAGGAGGAGGAGG 600
DB 1501 ACTTAAATGTGAAGAAAAAAGAGGGGGAAGAGGAGGAGGAGGAGGAGGAGGAGG 1560
QY 601 GGAAGACGCTCTTATGCTATGCTACTCTAATCTTTGCTCTACCTTCCACTTGG 660
DB 1561 GGAAGACGCTCTTATGCTATGCTACTCTAATCTTTGCTCTACCTTCCACTTGG 1620
QY 661 TTCTTAAATGAAGACTGTATGAAAGCTGGAGCCAGGCGAGGATGAGAGAGTGTG 720
DB 1621 TTCTTAAATGAAGACTGTATGAAAGCTGGAGCCAGGCGAGGATGAGAGAGTGTG 1680
QY 721 GTGTGTGTGGGGGGGGTGGGAGCAAGACAGACCTTTAGACAGAGAGAGCTGTAG 780
DB 1681 GTGTGTGTGGGGGGGGTGGGAGCAAGACAGACCTTTAGACAGAGAGAGCTGTAG 1740
QY 781 AGATCAATGAGCTTCTTTGAGACCCCTAGTGTACAGGAAATGTTCTTAACAGGTAG 840
DB 1741 AGATCAATGAGCTTCTTTGAGACCCCTAGTGTACAGGAAATGTTCTTAACAGGTAG 1800
QY 841 TGT 900

DB 1801 TGT 1860
QY 901 GAACAAACACCTTGGGCTCTTTAACCCTCTATTTCTACGCCCTCAGATGAAGATA 960
DB 1861 GAACAAACACCTTGGGCTCTTTAACCCTCTATTTCTACGCCCTCAGATGAAGATA 1920
QY 961 TGT 1020
DB 1921 TGT 1980
QY 1021 TAAACATTAACCTGAAGTGGGCTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 1080
DB 1981 TAAACATTAACCTGAAGTGGGCTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 2040
QY 1081 TGT 1140
DB 2041 TGT 2100
QY 1141 CAATGAAAGGCAACACCTCCGCTCTTACCCGCTTTGTGTGTGTGTGTGTGTGTGT 1200
DB 2101 CAATGAAAGGCAACACCTCCGCTCTTACCCGCTTTGTGTGTGTGTGTGTGTGTGT 2160
QY 1201 ACTCACTCTGGCACTGAACAGGCTTTTCTGACAGCTGGGGCTGTGAGGTGTGT 1260
DB 2161 ACTCACTCTGGCACTGAACAGGCTTTTCTGACAGCTGGGGCTGTGAGGTGTGT 2220
QY 1261 CTGAGACAGGAAACCTATCTTTTACTATGACATATTAACACAGGAGCTGTGAT 1320
DB 2221 CTGAGACAGGAAACCTATCTTTTACTATGACATATTAACACAGGAGCTGTGAT 2280
QY 1321 AGCCCTGAGCTGTGGCTTTTAAAGTTTGAATGATCTTGAAGAAATGTGTGTGT 1380
DB 2281 AGCCCTGAGCTGTGGCTTTTAAAGTTTGAATGATCTTGAAGAAATGTGTGTGT 2440
QY 1381 AGAAGCTTACCTGGCATGGGGCCCATCTCTCCAGCATTCGAATCTCAATCTGTG 1440
DB 2441 AGAAGCTTACCTGGCATGGGGCCCATCTCTCCAGCATTCGAATCTCAATCTGTG 2500
QY 1441 GTGCGTAAAGATTAATCTCTGGAATGTGAACATGTCTTCTTCTTCTGGGCTTGC 1500
DB 2501 GTGCGTAAAGATTAATCTCTGGAATGTGAACATGTCTTCTTCTTCTGGGCTTGC 2560
QY 1501 TGAAGAGTCCCAAGGCGAGCTGTGGGGGTAGGTGTGGGAAAAATCTCCAGA 1560
DB 2561 TGAAGAGTCCCAAGGCGAGCTGTGGGGGTAGGTGTGGGAAAAATCTCCAGA 2620
QY 1621 ACTCAGACATGTGGGCGGCTGTGAAGAGACTTTGTCAACA 1660
DB 2621 ACTCAGACATGTGGGCGGCTGTGAAGAGACTTTGTCAACA 2720
QY 1721 TTTTAAAGTGTGAAGAAAAAAGAGGGGGAAGAGGAGGAGGAGGAGGAGGAGGAGG 1780
DB 2721 TTTTAAAGTGTGAAGAAAAAAGAGGGGGAAGAGGAGGAGGAGGAGGAGGAGGAGG 2840
QY 1841 TGT 2900

RESULT 3
AC016464 207160 bp DNA linear HTG 01-MAR-2000
LOCUS AC016464
DEFINITION Mus musculus chromosome 11 clone RP23-40921 map 11, WORKING DRAFT
ACCESSION AC016464
VERSION AC016464.3 GI:7137319
KEYWORDS HTG; HTGS_PHASE1; HTGS_DRAFT.
SOURCE house mouse.
ORGANISM Mus musculus.
REFERENCE 1 (bases 1 to 207160)
AUTHORS Birren, B., Linton, L., Nusbaum, C. and Lander, E.
TITLE Mus musculus chromosome 11, clone RP23-40921
JOURNAL Unpublished
REFERENCE 2 (bases 1 to 207160)

Query Match 96.5%; Score 1601.2; DB 2; Length 207160;
 Best Local Similarity 99.0%; Pred. No. 0;
 Matches 161; Conservative 3; Mismatches 6; Indels 8; Gaps 4;

QY 1 AGGTTTACCCGATTTTGACACTAGATGGCATCCGTCACAGGGGTAGAGGTCAATGAAGCT 60
 |||||||
 Db 27190 AGGTTTACCCGATTTTGACACTAGATGGCATCCGTCACAGGGGTAGAGGTCAATGAAGCT 27249

QY 61 GACCAAGGCAAGTCCTTTCAGGGGGAAGAAATCAGG--AAAAAATTTTGAAGC 118
 |||||||
 Db 27250 GACCAAGGCAAGTCCTTTCAGGGGGAAGAAATCAGGAAAAAATTTTGAAGC 27309

QY 119 ATTTCAGGAAGCAAGATGAAATATTTGTACAAAACAGGTCTTCTCCACCATCGA 178
 |||||||
 Db 27310 ATTTCAGGAAGCAAGATGAAATATTTGTACAAAACAGGTCTTCTCCACCATCGA 27369

QY 179 CCGGGAGCT--CCACGATATGGACAGAAATGCTTACAGCTACATTCAAA--ACACAC 234
 |||||||
 Db 27370 CCGGGAGCTCCACGATATGGACAGAAATGCTTACAGCTACATTCAAAACACACAC 27429

QY 235 ACATCTTTTCCCTCC 294
 |||||||
 Db 27430 ACACACACACACACACACACACACACACACACACACACATGTTTCTTCCCTCC 27489

QY 295 CTCACATCTCCATCTCTCTGTGTGTCCTCCAAAGATGACATATGACTGTAGAAATCA 354
 |||||||
 Db 27490 CTCACATCTCCATCTCTGTGTGTCCTCCAAAGATGACATATGACTGTAGAAATCA 27549

QY 355 CACACACATTAAGCCCATCTGGGAGCCATTCAGACTGATCTTTTATCATTAAGGT 414
 |||||||
 Db 27550 CACACACATTAAGCCCATCTGGGAGCCATTCAGACTGATCTTTTATCATTAAGGT 27609

QY 415 TGAATCTTGGCAGCTGTGGGTTTAAAGTTTAAAGATTTTATCTAAGGCGACTAC 474
 |||||||
 Db 27610 TGAATCTTGGCAGCTGTGGGTTTAAAGTTTAAAGATTTTATCTAAGGCGACTAC 27669

QY 475 CTGCTCTCCCTGATATGTCGAATTCACCTGGCTTGGTACGCTAATGAAATGATCTAT 534
 |||||||
 Db 27670 CTGCTCTCCCTGATATGTCGAATTCACCTGGCTTGGTACGCTAATGAAATGATCTAT 27729

QY 535 GGTTCGACTTAATGTGAAGGAAAAAAGAAAGGAGGAAAAAGAGGAGGAGAAAGAG 594
 |||||||
 Db 27730 GGTTCGACTTAATGTGAAGGAAAAAAGAAAGGAGGAAAAAGAGGAGGAGAAAGAG 27789

QY 595 GGGAGGAAAACTGCTTTTATGCTTATGCTTACTTAACTTTTGTCTCACCCTTC 654
 |||||||
 Db 27790 GGGAGGAAAACTGCTTTTATGCTTATGCTTACTTAACTTTTGTCTCACCCTTC 27849

QY 655 ACTTGCTCTTCAATGGAAGACGTGATGAAGGCGGAGGAGCAGCAGGATAGAGAGA 714
 |||||||
 Db 27850 ACTTGCTCTTCAATGGAAGACGTGATGAAGGCGGAGGAGCAGCAGGATAGAGAGA 27909

QY 715 --GTGTGTGTGTGTGGGGGGGGGTGGGAGCAGCAGAGCAGGCTTATAGACAGAGAG 772
 |||||||
 Db 27910 GTGTGTGTGTGTGTGGGGGGGGGTGGGAGCAGCAGAGCAGGCTTATAGACAGAGAG 27969

QY 773 CCTGCTAGATCATGAGCTTCTTGTGAACCCCTAGTCTAAGAGATAGTCTCTAAC 832
 |||||||
 Db 27970 CCTGCTAGATCATGAGCTTCTTGTGAACCCCTAGTCTAAGAGATAGTCTCTAAC 28029

QY 833 CAGGTAGCTGTGTCAGTGTGCTGCTGGAAGSCTGCTGCTTGTCTTTTGTCTGCTG 892
 |||||||
 Db 28030 CAGGTAGCTGTGTCAGTGTGCTGCTGGAAGSCTGCTGCTTGTCTTTTGTCTGCTG 28089

QY 893 GCAGCCTTGAACAAACACCTGGCTCTTTGAACCCCACTATTCTCAGCCCTCAATGA 952
 |||||||
 Db 28090 GCAGCCTTGAACAAACACCTGGCTCTTTGAACCCCACTATTCTCAGCCCTCAATGA 28149

QY 953 AGAAGTAATGTATCTTGAAGATGATGATGATGATGATGATGATGATGATGATGATG 1012
 |||||||
 Db 28150 AGAAGTAATGTATCTTGAAGATGATGATGATGATGATGATGATGATGATGATGATG 28209

QY 1013 AGGTTTGTAAACATAAATGAAGTGGGTGTGGTGTAGTAAAGTAGCATGAATACATA 1072
 |||||||
 Db 28210 AGGTTTGTAAACATAAATGAAGTGGGTGTGGTGTAGTAAAGTAGCATGAATACATA 28269

QY 1073 AAAATATCTGTCAAGTGGCCAGAGCATCACTGTGTTAGAAACAAAGGCCCACTCGAA 1132
 |||||||
 Db 28270 AAAATATCTGTCAAGTGGCCAGAGCATCACTGTGTTAGAAACAAAGGCCCACTCGAA 28329

QY 1133 CAGCGGACATTTGAAGGACCAACCTCCGTGCTTACCCGTTGTTTGAACGGTG 1192
 |||||||
 Db 28330 CAGCGGACATTTGAAGGACCAACCTCCGTGCTTACCCGTTGTTTGAACGGTG 28389

QY 1193 TAAACGCAACTCACTCTGCGCACTGAACAGCTTTTGTCTGCAAGCTGGGCTCGAGG 1252
 |||||||
 Db 28390 TAAACGCAACTCACTCTGCGCACTGAACAGCTTTTGTCTGCAAGCTGGGCTCGAGG 28449

QY 1253 TGTGTCTCTGAGACAGAAACATCTTGTGTACTATGCAATGATAGTAAACACGAGC 1312
 |||||||
 Db 28450 TGTGTCTCTGAGACAGAAACATCTTGTGTACTATGCAATGATAGTAAACACGAGC 28509

QY 1313 TCTGATAGCCCTGAGCTGTGCTGCTTGAAGAAAGTTGATGCTTTGAAGAAATCGT 1372
 |||||||
 Db 28510 TCTGATAGCCCTGAGCTGTGCTGCTTGAAGAAAGTTGATGCTTTGAAGAAATCGT 28569

QY 1373 GGCCTTAAAGAGCCCTACCTGGCATGGGGCCATCTCTCCAGCATCCGAATCTCAAT 1432
 |||||||
 Db 28570 GGCCTTAAAGAGCCCTACCTGGCATGGGGCCATCTCTCCAGCATCCGAATCTCAAT 28629

QY 1433 CTGGTCTGTGCTGAAGATGAATCTCGGAATGATGATGATGATGATGATGATGATG 1492
 |||||||
 Db 28630 CTGGTCTGTGCTGAAGATGAATCTCGGAATGATGATGATGATGATGATGATGATG 28689

QY 1493 GGGCTTGTGAGAGAGTCCAGAGCTGTAGACGCTTGTGGGGGTAGTCTGGGAAAAATC 1552
 |||||||
 Db 28690 GGGCTTGTGAGAGAGTCCAGAGCTGTAGACGCTTGTGGGGGTAGTCTGGGAAAAATC 28749

QY 1553 TCCCAAGATTTTGAAGAGGCGCGCGGGGATGCAAAATCTTGAGATTCGGATCGCT 1612
 |||||||
 Db 28750 TCCCAAGATTTTGAAGAGGCGCGCGGGGATGCAAAATCTTGAGATTCGGATCGCT 28809

QY 1613 GTAGAGCAACTCAGACAGTCCGCGGCTGAAGAGCACTTGTCAAAACA 1660
 |||||||
 Db 28810 GTAGAGCAACTCAGACAGTCCGCGGCTGAAGAGCACTTGTGTCAAAACA 28857

RESULT 4
 AL646097/211173 bp DNA linear HTG 10-JAN-2002
 LOCUS Mus musculus chromosome 11 clone RP23-338M9, *** SEQUENCING IN
 DEFINITION
 ACCESSION AL646097 GI:18135259
 VERSION AL646097.5
 KEYWORDS HTG; HTGS_PHASE1; HTGS_DRAFT.
 SOURCE house mouse.
 ORGANISM Mus musculus
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
 1 (sites)
 McIlroy, K.
 DIRECT SUBMISSION
 Submitted (08-JAN-2002) Wellcome Trust Sanger Institute, Hinxton,
 Cambridgeshire, CB10 1SA, UK. E-mail enquiries:
 humquerry@sanger.ac.uk
 On Jan 11, 2002 this sequence version replaced gi:18072579.
 ----- Genome Center
 Center: Wellcome Trust Sanger Institute
 Center code: SC
 Web site: http://www.sanger.ac.uk
 Contact: humquerry@sanger.ac.uk
 Project Information
 Center project name: BM318M9
 ----- Summary Statistics
 Assembly program: XGAP4; version 4.5

|||||
Db 132882 AGAAGTAATGTAACCTTGGAGGATGATGGGTCAAGTAAGTAAGGAGAGGTTGGA 132823
OY 1013 AGGTTTGTACCATTAACAGAGTGGGGTGTGTAGTAAGTACCATTAATACATA 1072
Db 132822 AGGTTTGTACCATTAACAGAGTGGGGTGTGTAGTAAGTACCATTAATACATA 132763
OY 1073 AAAATATCTGTACAGTGGCCAGACATCTGTGTTCAGAACACACAGGCCACTGAAA 1132
Db 132762 AAAATATCTGTACAGTGGCCAGACATCTGTGTTCAGAACACACAGGCCACTGAAA 132703
OY 1133 CAGGGGACATTTGAAAAGGACCAACCTCGTGTCTTCTACCCGTTTGTGTACCGTG 1192
Db 132702 CAGGTGGACATTTGAAAAGGACCAACCTCGTGTCTTCTACCCGTTTGTGTACCGTG 132643
OY 1193 TAAACGCACTCACTCTCTGCGACATGAACAGGCTTTTCTGTCAGACACTGGGGTCTGAGG 1252
Db 132642 TAAACGCACTCACTCTCTGCGACATGAACAGGCTTTTCTGTCAGACACTGGGGTCTGAGG 132583
OY 1253 TGTGTCTCTGAGACAGGAAAACATCTGTGTACTATGCGATAGTATACCGAGAGC 1312
Db 132582 TGTGTCTCTGAGACAGGAAAACATCTGTGTACTATGCGATAGTATACCGAGAGC 132523
OY 1313 TCTGAGTATACCCCTGAGCTGTGGTCCCTTTAGAAAAGTTTGATGCTTGAAGAAATCGT 1372
Db 132522 TCTGAGTATACCCCTGAGCTGTGGTCCCTTTAGAAAAGTTTGATGCTTGAAGAAATCGT 132463
OY 1373 GGGCTTAAAGAAAGCCCTACCTGGCATGGGGGCCCATCTCTCCAGCACTCCGAAATCTCAAT 1432
Db 132462 GGGCTTAAAGAAAGCCCTACCTGGCATGGGGGCCCATCTCTCCAGCACTCCGAAATCTCAAT 132403
OY 1433 CTGCTCTGTGCGTAAAGATTAAGATCTCTCGGAATGTATACCATCTGCTTTCTTCTCT 1492
Db 132402 CTGCTCTGTGCGTAAAGATTAAGATCTCTCGGAATGTATACCATCTGCTTTCTTCTCT 132343
OY 1493 GGGCTTCTGAGGAAGTCCCGAGCGATAGACCTCTGTGGGGTAGCTCTGGGAAAAATC 1552
Db 132342 GGGCTTCTGAGGAAGTCCCGAGCGATAGACCTCTGTGGGGTAGCTCTGGGAAAAATC 132283
OY 1553 TCCCAAGATTTTAAAGAGGGGCGGGGGGATGAGAAAACCTTGAGATTCGGTAGTCGCT 1612
Db 132282 TCCCAAGATTTTAAAGAGGGGCGGGGGGATGAGAAAACCTTGAGATTCGGTAGTCGCT 132223
OY 1613 GTAGAGCACTCAGACAGTGGGGCCCTGAGAGGACTTGTGCAACA 1660
Db 132222 GTAGAGCACTCAGACAGTGGGGCCCTGAGAGGACTTGTGCAACA 132175

RESULT 5
AX354812 493 bp DNA linear PAT 06-FEB-2002
LOCUS AX354812
DEFINITION Sequence 6 from Patent W00179290.
ACCESSION AX354812
VERSION AX354812.1 GI:18619543
KEYWORDS
SOURCE Mus sp.
ORGANISM Mus sp.
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
REFERENCE 1 (sites)
AUTHORS Drucker,D.J. and Lowshin,J.A.
TITLE GIP-2 receptor gene promoter and uses thereof
JOURNAL Patent: WO 0179290-A 6 25-OCT-2001;
1149336 ONTARIO INC. (CA)
Location/Qualifiers
FEATURES
source 1..493
/organism="Mus sp."
/db_xref="taxon:10095"
401..>466
/note="unnamed protein product"
/codon_start=1
/protein_id="CAD22992.1"
/db_xref="GI:18619544"

BASE COUNT 100 a 132 c 155 g 104 t 2 others
ORIGIN
Query Match 10.5%; Score 175; DB 6; Length 493;
Best Local Similarity 99.5%; Pred. No. 5.7e-42;
Matches 186; Conservative 0; Mismatches 0; Indels 1; Gaps 1;
OY 1474 ATGCTTGTCTTTCTTCTGCGGCTTGTGAGGAGTCCCGAGGCGGATGAGAACTT 1533
Db 1 ATGCTTGTCTTTCTTCTGCGGCTTGTGAGGAGTCCCGAGGCGGATGAGAACTT 60
OY 1534 GGTAGTCTGGGAAAATCTCCCAAGATTTTAAAGAGGGGCGGGGATGAGAACTT 1593
Db 61 GGTAGTCTGGGAAAATCTCCCAAGATTTTAAAGAGGGGCGGGGATGAGAACTT 119
OY 1594 GGAGATTCGGTAGATCGCTGTAGAGCACTCAGACAGTGGGGCTGAAAGAGACTTGT 1653
Db 120 GGAGATTCGGTAGATCGCTGTAGAGCACTCAGACAGTGGGGCTGAAAGAGACTTGT 179
OY 1654 GCAACA 1660
Db 180 GCAACA 186

RESULT 6
AC026591/c 198271 bp DNA linear HTG 27-AUG-2001
LOCUS AC026591/c
DEFINITION Homo sapiens chromosome 17 clone RP11-655D3 map 17, *** SEQUENCING
IN PROGRESS ***. 4 ordered pieces.
ACCESSION AC026591
VERSION AC026591.9 GI:15291083
KEYWORDS HTG: HTGS_PHASE2; HTGS_FULLTOP; HTGS_ACTIVEFIN.
SOURCE human
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
REFERENCE 1 (bases 1 to 198271)
AUTHORS Birren,B., Linton,L., Nusbaum,C. and Lander,E.
JOURNAL Homo sapiens chromosome 17, clone RP11-655D3
TITLE Unpublished
JOURNAL 2 (bases 1 to 198271)
AUTHORS Birren,B., Linton,L., Nusbaum,C., Lander,E., Abraham,H., Allen,N.,
Anderson,S., Baldwin,J., Barna,N., Bastien,V., Beda,F.,
Boguslavsky,L., Boukhalter,B., Brown,A., Burkett,G.,
Campobiano,A., Castle,A., Choepel,Y., Colangelo,M., Collins,S.,
Collinsmore,A., Cooke,P., Dearrellano,K., Dewar,K., Diaz,J.S.,
Dodgson,S., Domingo,M., Doyle,M., Ferreira,P., FitzHugh,W., Gage,D.,
Galagan,J., Gardyna,S., Glende,S., Goyette,M., Graham,L.,
Grand-Pierre,N., Grant,G., Hagos,B., Harford,A., Horton,L.,
Howland,J.C., Iliev,I., Johnson,R., Jones,C., Kann,L., Karatas,A.,
Klein,J., Labocque,K., Lamazares,R., Landers,T., Lehoczy,J.,
Levine,R., Liu,C., Locke,K., Macdonald,P., McPheters,R.,
McCarthy,R., McEwan,P., McCurt,A., Miranda,C., Mlenga,V., Morrow,J.,
Meldrum,T., Meneses,L., Milnova,T., Miranova,C.H., O'Connor,I., O'Donnell,P.,
Murphy,T., Naylor,J., Norman,C.H., O'Connor,I., O'Donnell,P.,
O'Neil,D., Oliver,T.M., Oliver,J., Peterson,K., Pierre,N.,
Pisani,C., Pollara,V., Raymond,C., Riley,R., Rogov,P., Rothman,D.,
Roy,A., Santos,R., Schauer,S., Severy,S., Subramanian,A., Talamas,J.,
Stange-Thomann,N., Stojanovic,N., Tirrelli,A., Travers,M., Trigilio,J.,
Vassiliev,H., Viel,R., Vo,A., Wilson,B., Wu,X., Wyman,D., Ye,W.J.,
Young,G., Zainoun,J., Zimmer,A. and Zody,M.
Direct Submission
Submitted (22-MAR-2000) Whitehead Institute/MIT Center for Genome
Research, 320 Charles Street, Cambridge, MA 02141, USA
On Aug 26, 2001 this sequence version replaced gi:14547870.
All repeats were identified using RepeatMasker:
Smit,A.F.A. & Green,P. (1996-1997)
http://ftp.genome.washington.edu/RM/RepeatMasker.html
----- Genome Center
Center: Whitehead Institute/ MIT Center for Genome Research
Center code: WIBR

Db 144074 CAACA 144069

|||||

RESULT 7

AC087646/c

LOCUS

Homo sapiens chromosome 17 clone RP11-773021 map 17, WORKING DRAFT
SEQUENCE, 12 unordered pieces.

ACCESSION

AC087646 GI:14269760

VERSION

HTG: HTGS_PHASE1; HTGS_DRAFT; HTGS_FULLTOP.

SOURCE

human

ORGANISM

Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Homidae; Homo.

REFERENCE

1 (bases 1 to 210719)
Birren, B., Linton, L., Nusbaum, C. and Lander, E.

TITLE

Homo sapiens chromosome 17, clone RP11-773021

JOURNAL

Unpublished

AUTHORS

2 (bases 1 to 210719)
Birren, B., Linton, L., Nusbaum, C., Lander, E., Allen, N., Anderson, S.,
Barra, N., Bastien, V., Boguslavsky, L., Boukhgalter, B., Brown, A.,
Cammarata, J., Campiano, A., Choquet, Y., Colangelo, M., Collins, S.,
Collamore, A., Cooke, P., Dearellano, K., Dewar, K., Diaz, J.S.,
Dodge, S., Faro, S., Ferreira, P., Fitzhugh, W., Gage, D., Galagan, J.,
Gardyna, S., Ginde, S., Goyette, M., Graham, L., Grand-Pierre, N.,
Hagos, B., Heath, A., Horton, L., Hulme, W., Iliev, I., Johnson, R.,
Jones, C., Karas, A., Larocque, K., Lamazares, R., Landers, T.,
Lehoczky, J., Levine, R., Liu, G., Maclean, C., Macdonald, P.,
Marquis, N., Matthews, C., McCarthy, M., McEwan, P., McKernan, K.,
McPheters, R., Meldrum, J., Meneus, L., Minova, T., Mlenga, V.,
Murphy, T., Naylor, J., Nguyen, C., Norbu, C., Norman, C.H.,
O'Connor, T., O'Donnell, P., O'Neill, D., Oliver, J., Peterson, K.,
Phunkhang, P., Pierre, N., Pollara, V., Raymond, C., Retta, R.,
Rieback, M., Riley, R., Rise, C., Rogov, P., Roman, D., Rosetti, M.,
Roy, A., Santos, R., Schauer, S., Schupbach, R., Seaman, S., Severy, P.,
Souarez, C., Spencer, B., Stange-Thomann, N., Stojanovic, N.,
Strauss, N., Subramanian, A., Talamas, J., Testfay, S., Theodore, J.,
Travers, M., Travis, N., Triggillo, J., Vassiliev, H., Viel, R., Vo, A.,
Wilson, B., Wu, X., Wyman, D., Ye, W.J., Young, G., Zainoun, J.,
Zemek, L., Zimmer, A. and Zody, M.

COMMENT

Submitted (15-JAN-2001) Whitehead Institute/MIT Center for Genome
Research, 320 Charles Street, Cambridge, MA 02141, USA
On May 31, 2001 this sequence version replaced gi:13493095.
All repeats were identified using RepeatMasker:
http://ftp.genome.washington.edu/RM/RepeatMasker.html

TITLE

Genome Center

JOURNAL

Center: Whitehead Institute/ MIT Center for Genome Research

AUTHORS

Center code: WIBR

REFERENCE

Web site: http://www-seq.wi.mit.edu

TITLE

Contact: sequence_submissions@genome.wi.mit.edu

JOURNAL

Project Information

AUTHORS

Center project name: L11995

REFERENCE

Center clone name: 773-O-21

TITLE

Summary Statistics

JOURNAL

Sequencing vector: Plasmid; n/a; 100% of reads

AUTHORS

Chemistry: Dye-terminator Big Dye; 100% of reads

REFERENCE

Assembly program: Phrap; version 0.960731

TITLE

Consensus quality: 205962 bases at least Q40

JOURNAL

Consensus quality: 208431 bases at least Q30

AUTHORS

Consensus quality: 209109 bases at least Q20

REFERENCE

Insert size: 205000; agarose-1p

TITLE

Quality coverage: 8.4 in Q20 bases; sum-of-contigs

JOURNAL

Quality coverage: 8.2 in Q20 bases; sum-of-contigs

AUTHORS

* NOTE: This is a 'working draft' sequence. It currently

REFERENCE

* consists of 12 contigs. The true order of the pieces

TITLE

* is not known and their order in this sequence record is

JOURNAL

* arbitrary. Gaps between the contigs are represented as

* runs of N, but the exact sizes of the gaps are unknown.
* This record will be updated with the finished sequence
* as soon as it is available and the accession number will
* be preserved.
1 75254: contig of 75254 bp in length
75255 75354: gap of 100 bp
75355 76881: contig of 1527 bp in length
76882 76981: gap of 100 bp
76982 78538: contig of 1557 bp in length
78539 78638: gap of 100 bp
78639 79946: contig of 1308 bp in length
79947 80046: gap of 100 bp
80047 82098: contig of 2052 bp in length
82099 82198: gap of 100 bp
82199 86309: contig of 4111 bp in length
86310 86409: gap of 100 bp
86410 89601: contig of 3192 bp in length
89602 89701: gap of 100 bp
89702 105000: contig of 15299 bp in length
105001 105100: gap of 100 bp
105101 121101: contig of 16001 bp in length
121102 121201: gap of 100 bp
121202 154071: contig of 32870 bp in length
154072 154171: gap of 100 bp
154172 193985: contig of 39814 bp in length
193986 194085: gap of 100 bp
194086 210719: contig of 16634 bp in length.
Location/Qualifiers
1. 210719
/organism="Homo sapiens"
/db_xref="taxon:9606"
/chromosome="17"
/map="17"
/clone="RP11-773021"
/clone_1b="RP11-773021 Human Male BAC"
1. 75254
/note="assembly-fragment"
clone_end:SP6
vector_side:left
75355. 76881
/note="assembly-fragment"
76882. 78538
/note="assembly-fragment"
78639. 79946
/note="assembly-fragment"
80047. 82098
/note="assembly-fragment"
82199. 86309
/note="assembly-fragment"
86410. 89601
/note="assembly-fragment"
89702. 105000
/note="assembly-fragment"
105101. 121101
/note="assembly-fragment"
121202. 154071
/note="assembly-fragment"
154172. 193985
/note="assembly-fragment"
194086. 210719
/note="assembly-fragment"
clone_end:T7
vector_side:right

BASE COUNT 58923 a 47911 c 46884 g 55896 t 1105 others

Query Match 6.1%, Score 101.6; DB 2; Length 210719;
Best Local Similarity 49.4%; Pred. No. 1.9e-19;
Matches 803; Conservative 2; Mismatches 696; Indels 125; Gaps 16;

07 115 AAGCATTCAGAGCAAGCAGATGATATTTGACAAACAGGTGCTTCCGCCACCAT 174
||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 73808 AAAAGTGCAGAGCATGAGATGATATTTCTTGCAGAGCAGATTTCTTAAGACAG 73749

QY 175 GCGACCCGGAGCTCCAGTATGAGACAGATAGCTTTACAGCTTACATTCATAAAGACAC 234
 Db 73748 ACCATTTTCCAGC-----ATGCACTTGGAGATGATCTTTAATGAGCACTCCACAGAGCG 73693
 QY 235 AC 294
 Db 73692 AGTTGACCCACATTCATGAC-----CAAGTCCATCCCTTGGGTACTTCTCTCCCTCG 73638
 QY 295 CTCACACTTCCTCCATCTCTGCTGCTGCTCCCAAGAGATACCATATGCTGTAGAAATCA 354
 Db 73637 GGTACACTCCATCTGCTGCTGCTGCTCCAGAGAGATGACACGTCATTCACAGAGGGA 73578
 QY 355 CACACACATTAAGAGCCATTCGAGAGCCATTCAGACTGATCTTTTATCATTAAGGT 414
 Db 73577 CACACACTGAGTCAATTCATTCAGAGCTTCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 73518
 QY 415 TGAATTCCTCCAGCTGTGGGTTTAAAGTTTAAAGTTTAAAGTTTAAAGTTTAAAGTT 474
 Db 73517 TAAATTCCTGACATGAGATTTTAAAGTTTAAAGTTTAAAGTTTAAAGTTTAAAGTT 73458
 QY 475 CTGCTTCCTGTCATGATGTCAGATTCAGTGGCTTGGTCCAGCTAATGCAATGATCTAT 534
 Db 73457 CTGCTTCCTGTCATGATGTCAGATTCAGTGGCTTGGTCCAGCTAATGCAATGATCTAT 73398
 QY 535 GGTTCACCTTAAATGTGAAAGAAAAAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 594
 Db 73397 TGTTCACCTTAAATGTGAAAGAAAAAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 73350
 QY 595 GGGAG 654
 Db 73359 AC 73300
 QY 655 ACTTGTCTTCAATGAGAAAGACTGATAGAAAGCTGGAGCCAGCCAGGATGAGAGGA 714
 Db 73299 ACCAATCAAGACATCTTCCACCTGGGCTCCCAATTTGGAGGCTGGAGACATTCAGGG 73240
 QY 715 GTGTGTGTGTGTGG 774
 Db 73239 AGCGTGGAGCTGGCTGGGAGAGGTGAGTAAAGATTCC-----AAGCAGAGCTTTGG 73184
 QY 775 TCCAGAGACATGAGCTTCTTGTAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 827
 Db 73183 GAGGATGGCGATGAGCTGCTTAAGAACTACCGGAGAGAGAGAGAGAGAGAGAGAGAG 73124
 QY 828 CTAACGAGTACCTGTGTGACGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 879
 Db 73123 CTTCCCAAGTACCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 73064
 QY 880 TTTTGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 939
 Db 73063 GTTTCCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 73004
 QY 940 ACCCTCAGATGAAGAGTAAATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 998
 Db 73003 AGCTGTCAAAATGAGAGATTCATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 72944
 QY 999 GGGCAGAGAGGCTGGAAGTTTGTAAACCAATGAAGTGGGCTGTTGTTAGTAAGTAG 1058
 Db 72943 GGAATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 72888
 QY 1059 CCATGAATACCATTAATAATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1118
 Db 72887 GCTGCTTGAATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 72828
 QY 1119 GCGCCCACTGACAGACAGGAGCAATTTGAAGGACACACCTCGCTTCTCTACACCGTT 1178
 Db 72827 ACATGTGCTTGAAGACAGGAGGCTTCAAGATGAGGAATGTTGCACTTTTCAACACT 72768
 QY 1179 GTTTGCTTACCGTGAAGCACTCACTGCGGACGTAACAGGCTTTGCTGCGAGAC 1238
 Db 72767 GTTTGCTTACCGTGAAGCACTCACTGCGGACGTAACAGGCTTTGCTGCGAGAC 72713

QY 1239 CTGGGCTCTGGAGCTGTTGCTCTGAGACAGAGAAACTCATCTTGTACTATGCGATAGT 1298
 Db 72712 CTTGGGCTGGAGCTGTTGCTCTGAGCTGAGACACACACTCTCTGCACTCTGCTGAT 72653
 QY 1299 AGTAACACAGAGCTCTGATATGAGCT-----GAGCTGGGCGCTTTAGAAAATT 1350
 Db 72652 AAACAGAGCCCAAGAGAGATTTTGCTTTGTTAGCTTCAGCCAGCTGGACAAAGCT 72593
 QY 1351 TGAATCTTGAAGAAAGT--CGTGGCTTAAAGAAAGCTTACCTGGCATGGGGCCCATC 1408
 Db 72592 TAATGCTTGAAGAAAGTACCTGCTCAAAAAGACACCTGCTTAACTGGCTGAG 72533
 QY 1409 CTCTCAGCATCCGATCTCATCTGCTGCTGCTA-----1447
 Db 72532 CCCATCGCTTTCGACAGCATCTCTTACTCTGACACCCAGCTGCTCAGCCAGC 72473
 QY 1448 -----AGATTAAGATTCCTGGAATGTAAC-----ATGCTGCTTTTCTCTG 1493
 Db 72472 TTGTCATTTGAGACACACACTTACAGGCTGAGCCCGCTGACCGCTTCTTCTCTC 72413
 QY 1494 GCGTGTGAGAGATTCACAGCAGCAGCTGATGCTTTGGGGTAAAGTCTGGAAAAATCT 1553
 Db 72412 AGCTGTCAAGGAAGTCCAGAAAGCAGAGCTGAT--AGGAAAGTCTGGAAAAATCT 72354
 QY 1554 CCCAAGATTTTGAAGAGGCGAGCGGGGAGTGA-----GAACTTGAGAG 1598
 Db 72353 CCTGCTTTTGGGGGGGAGGAGGCGGGGATGAGCCAGGCGCGAAGAACTCTGAAGA 72294
 QY 1599 TTGCGTACATGCTGCTGAGACAACTCAGACAGTC-----GGCGGCTGAAAGACTTGTG 1654
 Db 72293 CTCCGTAGATTTGCTGAGACCGGCTCAGACACTCTGGGCGACGCTGAGAGATTTGTG 72234
 QY 1655 CAACA 1660
 Db 72233 CAACA 72228

RESULT 8
 AC069006/c
 LOCUS
 DEFINITION Homo sapiens chromosome RPc1-11 clone RP11-773021, WORKING DRAFT
 SEQUENCE, 16 unordered pieces.
 AC069006
 AC069006.4 GI:11612641
 HTG: HTGS_PHASE1; HTGS_DRAFT.
 SOURCE human.
 ORGANISM Homo sapiens
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 Mammalia; Eutheria; Primates; Catarrhini; Homiidae; Homo.
 1 (bases 1 to 213359)
 Waterston,R.H.
 The sequence of Homo sapiens clone
 Unpublished
 2 (bases 1 to 213359)
 Waterston,R.H.
 Direct Submission
 Submitted (16-MAY-2000) Genome Sequencing Center, Washington
 University School of Medicine, 4444 Forest Park Parkway, St. Louis,
 MO 63108, USA
 On Dec 11, 2000 this sequence version replaced gi:9280811.

----- Genome Center -----
 Center: Washington University Genome Sequencing Center
 Center code: WUGSC
 Web site: http://genome.wustl.edu/gsc/index.shtml
 ----- Project Information -----
 Center project name: H-NH073021
 ----- Summary Statistics -----
 Sequencing vector: M13: 100%
 Sequencing vector: plasmid: 0%
 Chemistry: Dye-primer ET; 100% of reads
 Chemistry: Dye-terminator Big Dye; 0% of reads
 Assembly program: Phrap; version 0.990319

Db	84362	CCTCCCAAGTACCTTGGCTCAAGCAACTGTGGTCAAGAAAGACCTGCCATTACGGTTCTGGT	84303
Qy	880	TTTGTGCTGCTGTGTGACGCTTGACAAACACCCTCGCTCTTTGAACCCCACTATTCTTC	939
Db	84302	GTTTACCTGCTGTGTGGCTTTGGGGAGACGCATTCACACCTCTGTGAGCCCAATTGTTCTTC	84243
Qy	940	AGCCCTCAGATGAAAGCAATGTGA-TCTTGAGAGATACATGAGGTTCAAGTGAACTA	998
Db	84242	AGCTGTCAAAATGAGAAATTCAATCTGATATCTCTCATGGGTTATTGTCAAGATGTGCCAAATA	84183
Qy	999	GGCGAGAGGGTGGAGGTTTGTGTAACCATTAACCTGAAGTGGGGTGTGGTTAGTAAGTAG	1058
Db	84182	GGATGTGGGTTAAAAAGAGCT---TGTCAAATGTGAGTGTGGGGACACCTCACTGCTCTGT	84127
Qy	1059	CCATGATATACCTTAAAAATATCTGTCAAGTGGCCAGAGCATCTAGTGTGTTCAAGAACAA	1118
Db	84126	GCTGTGTAGTAAGTATGATATCATCATCAATTCCATTTCAAAAGGGTCTGTCAAGTGGGCTGAGCC	84067
Qy	1119	CGGGCCACTCGAAGACACGGGGCAATTGAAAGGACCAACCCGCGTGTCCATACCCTGT	1178
Db	84066	ACATGTGCTTTAGAACAAAGTGGGCATTTCAAGATGAGTAATGTTGCATTTTCAACCACT	84007
Qy	1179	GTTTGTGTTACCGTGTAAACGCAACTCAACTCTCGGCACATGAACAGGCTTTCTCGCAGAC	1238
Db	84006	GTTTGTGTAATGTGTAGGCAATACCTA-----CGCGCGTAATGGGTTTTCGTTCTATGG	83952
Qy	1239	CTGGGCTTTCGAGTGTGTCTCTGAGACAGAAACTCATTTGTTACTATGCGATAGT	1298
Db	83951	CCTGGCGTGGGAGCGCTGTTCTCTCTGAGACTGAGCAACCACTCTCGCAGTCTGCAAT	83892
Qy	1299	AGTAAACACGGAGCTCTGAGATAGGCCCT-----GAGCTGTGGCGTTTGAAGAAAGTT	1350
Db	83891	AAACAGAACCCAGAGAGATTTTGGCTTTGTGTAGCCTCAGCCCAAGCTTGACCAAGCT	83832
Qy	1351	TGATGCTTTAGAAAGAAATCGTGGCTTAAAGAG-----CCTACCTGGCATGGGG	1401
Db	83831	TAAATGCTTTAGAAAGAAATACCTCTGCTCAAAAGACACCTGCTCTAATACGTGCTCGAG	83772
Qy	1402	GCCCATCTCTCCACCAATCCGAATCTC-----AAT	1432
Db	83771	CCCCCTGCCCTTCTCTGACGCCATTCCTCTGTAGTCTGAACCCAGCGCTGTCCAGCCAGC	83712
Qy	1433	CTGGTCGTGTGGGTAGAAATAGAAATCTCTGGAATGGTAACCATGTCTTCTTTCTTCT	1492
Db	83711	TTTCGTGCAATTGAGCGAAACCTTCAGGGTGGAGCCCGCTTGACCCGCTTGTCTTCTCTCT	83652
Qy	1493	GGGCTTGCCTGAGAGATGCCAGCAGCGTAGACGTCTTGGGGGTAGGCTGTGGGAAAAATTC	1552
Db	83651	CAGCTCTGCMAAGGAAGTCCCAAGAAAGCACAGCTGACTT-AGGGAAAGGTCTGGGAAAAAATTC	83593
Qy	1553	TCCCAAGATTTTAGAGAGGGCAGGGGGGGATGA-----GAAACTTGGAG	1597
Db	83592	TCCCTGCTTTTGGGGGGGAGGGGGGGGAGATGAGCCAGGGCCGGAAGGAACCTCTGAAG	83533
Qy	1598	ATTGGGTAGATGCTGTAGAGCACTCAGACAGTC-----GGCGGCTGGAAGAGACTGT	1653
Db	83532	ACTCGGTGAATGTGCTGTAGACCGCTCAGACACTCTCGCGCGAGAGGTGAGAGGATTTGT	83473
Qy	1654	GCAACA 1660	
Db	83472	GCAACA 83466	

[illegible]

REFERENCE
AUTHORS
TITLE
JOURNAL
REFERENCE
AUTHORS

Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Hominiidae; Homo. 1 (bases 1 to 49665)
Birren, B., Linton, J., Nusbaum, C. and Landey, E.
Homo sapiens chromosome 15, clone RP11-615H7

2 (bases 1 to 49665)

Birren, B., Linton, L., Nusbaum, C., Lander, E., All, A., Allen, N., Anderson, S., Barra, N., Bastien, V., Boguski, L., Boucknight, B., Brown, A., Camarata, J., Campiagno, A., Chang, J., Chazaro, B., Chopel, Y., Colangelo, M., Collins, S., Collymore, A., Cook, A., Cooke, P., Deavello, K., Dewar, K., Diaz, J. S., Dodge, S., Faro, S., Ferreira, P., Fitzhugh, W., Gage, D., Galagan, J., Gaidyna, S., Ginde, S., Gord, S., Goyette, M., Graham, L., Grand-Pierre, N., Hagos, B., Heaford, A., Horton, L., Hulme, W., Iliev, I., Johnson, R., Jones, C., Kamat, A., Karlas, A., Kells, C., Labocque, K., Lamazares, R., Landers, T., Lehoczy, J., Levine, R., Liu, G., Maclean, C., MacDonald, P., Major, J., Margulis, N., Matthews, C., McCarthy, M., McEwen, P., McKernan, K., McPeckers, R., Meldrum, J., Menus, L., Mltova, T., Menges, V., Murphy, T., Naylor, J., Nguyen, C., Norbu, C., Norman, C. H., O'Connor, T., O'Donnell, P., O'Neill, D., Oliver, J., Peterson, R., Phunkhang, P., Pierre, N., Pollara, V., Raymond, C., Retta, R., Riback, M., Riley, R., Rise, C., Rogov, P., Roman, J., Rosetti, M., Roy, A., Santos, R., Schauer, S., Schuback, R., Seaman, S., Severy, P., Spencer, B., Stange-Thomann, N., Stojanovic, N., Strauss, N., Subramanian, A., Talamas, J., Tesfaye, S., Theodore, J., Topham, K., Travers, M., Travis, N., Trifilio, J., Vassiliev, H., Viel, R., Vo, A., Wilson, B., Wu, X., Wyman, D., Ye, M. J., Young, G., Zaimou, J., Zambek, L., Zimmer, A., and Zody, M.

TITLE
JOURNAL
COMMENT

Submitted (22-NOV-2001) Whitehead Institute/MIT Center for Genome Research, 320 Charles Street, Cambridge, MA 02141, USA
All repeats were identified using RepeatMasker:
Smit, A.F.A. & Green, P. (1996-1997)
<http://ftp.genome.washington.edu/RM/RepeatMasker.html>

Center: Whitehead Institute/ MIT Center for Genome Research
Center code: WIBR
Web site: <http://www-seq.wi.mit.edu>
Contact: sequence_submissions@genome.wi.mit.edu
----- Project Information

* NOTE: This record contains 61 individual
* sequencing reads that have not been assembled into
* contigs. Runs of N are used to separate the reads
* and the order in which they appear is completely
* arbitrary. Low-pass sequence sampling is useful for
* identifying clones that may be gene-rich and allow
* overlap relationships among clones to be deduced.
* However, it should not be assumed that this clone
* will be sequenced to completion. In the event that
* the record is updated, the accession number will
* be preserved.

*	1	685:	contig of 685 bp in length
*	686	785:	gap of 100 bp
*	786	1524:	contig of 739 bp in length
*	1525	1624:	gap of 100 bp
*	1625	2380:	contig of 676 bp in length
*	2301	2400:	gap of 100 bp
*	2401	3073:	contig of 673 bp in length
*	3074	3173:	gap of 100 bp
*	3174	3870:	contig of 697 bp in length
*	3871	3970:	gap of 100 bp
*	3971	4667:	contig of 697 bp in length
*	4668	4767:	gap of 100 bp
*	4768	5508:	contig of 741 bp in length
*	5509	5608:	gap of 100 bp
*	5609	6387:	contig of 739 bp in length
*	6348	6447:	gap of 100 bp
*	6448	7197:	contig of 749 bp in length
*	7197	7296:	gap of 100 bp

KEYWORDS

HTG: HTGS-PHASE1.
Rattus norvegicus

SOURCE

Rattus norvegicus

ORGANISM

Rattus norvegicus

REFERENCE

1 (bases 1 to 102783)
Muzny,D.M., Adams,C., Adio-Oduola,B., Ali-osman,F.R., Allen,C., Alstbrooks,S.L., Amaralunga,H.C., Are,J.R., Banks,T., Barbata,J., Benton,J., Blmage,K., Blankenburg,K., Bonnin,D., Bouck,J., Bowler,S., Brieva,M., Brown,E., Brown,M., Bryant,N.P., Buhay,C., Burck,P., Burdett,C., Burrell,K.L., Byrd,N.C., Caron,T.F., Carter,M., Cavazos,S.R., Chacko,J., Chavez,D., Chen,G., Chen,R., Chen,Z., Chowdhury,I., Christopoulos,C., Cleveland,C.D., Cox,C., Coyle,M.D., Dathorne,S.R., David,R., Davila,M.L., Davis,C., Davy,A.L., Ding,Y., Dinh,H.H., Douthwaite,K.J., Drepper,H., Dugan-Rocha,S., Durbin,K.J., Earnhart,C., Edgar,D., Edwards,C.C., Elhaj,C., Escotto,M., Falls,T., Ferraguto,D., Flagg,N., Ford,J., Foster,P., Frantz,P., Gabisi,A., Gao,J., Garcia,A., Garner,T., Garza,N., Gill,R., Gorrell,J.H., Guevara,M., Gunaratne,P., Hale,S., Hamilton,K., Harris,C., Harris,K., Hart,M., Havlak,P., Hawes,A., Hernandez,J., Hernandez,O., Hodgson,A., Hogues,M., Hollway,C., Hollins,B., Homs,F., Howard,S., Huber,J., Hulik,S., Hume,J., Jackson,L.E., Jacobson,B., Jia,Y., Johnson,R., Jolivet,S., Joudah,S., Karlsson,E., Kelly,S., Khan,U., King,L., Korvah,J., Kovar,C., Kratovic,J., Kuresh,A., Landry,N., Leal,B., Lewis,L.C., Lewis,L., Li,J., Li,Z., Lichtarge,O., Lieu,C., Liu,J., Liu,W., Loulseged,H., Lozado,R.J., Lu,X., Lucier,A., Lucier,R., Luna,R., Ma,J., Maheshwari,M., Mapa,P., Martin,R., Martindale,A., Matinez,E., Massey,E., Mawhney,E., McLeod,M.P., Meador,M., Mel,G., Metzger,M., Miner,G., Miner,G., Mitchell,T., Mohabbat,K., Morgan,M., Morris,S., Moser,M., Neal,D., Newton,J., Newton,N., Nguyen,A., Nguyen,G., Otagunye,N., Oviedo,R., Pace,A., Payton,B., Ogden,M., Okunoye,G., Otagunye,N., Oviedo,R., Pace,A., Payton,B., Peery,J., Peters,L., Peters,L., Picken,R., Primus,E., Pu,L.L., Quiles,M., Ren,Y., Rives,M., Rojas,A., Rojibokan,I., Rolfe,M., Ruiz,S., Severy,G., Scherer,S., Scott,G., Shen,H., Shostkati,N., Slason,I., Sodergren,E., Sonalke,T., Sparks,A., Stanley,H., Stone,H., Sutton,A., Swalek,A., Tabor,P., Tamerisa,A., Tamerisa,K., Thomas,S., Usmani,K., Vasquez,L., Vera,V., Villalón,D., Vinson,R., Wallington,S., Williams,G., Williams,R., Washington,C., Worley,K., Wu,C., Wu,Y., Wu,Y.F., Zhou,J., Zorrilla,S., Nelson,D., Weinstein,G., and Gibbs,R.

TITLE

Unpublished
2 (bases 1 to 102783)
Submitted (17-SEP-2001) Human Genome Sequencing Center, Department of Molecular and Human Genetics, Baylor College of Medicine, One Baylor Plaza, Houston, TX 77030, USA
On Dec 20, 2001 this sequence version replaced g1:16901736.

COMMENT

Center: Baylor College of Medicine
Center code: BCM
Web site: <http://www.hgsc.bcm.tmc.edu/>
Contact: hgsc-help@bcm.tmc.edu
Project Information
Center project name: GEMB
Center clone name: CH230-11B19
Summary Statistics
Assembly program: Phrap; version 0.990329first call to findhaplist

Consensus quality: 92140 bases at least Q40
Consensus quality: 98027 bases at least Q30
Consensus quality: 102801 bases at least Q20
Estimated insert size: 79332; sum-of-contigs estimation
Quality coverage: 0x in Q20 bases; agarose-fp estimation
Quality coverage: 1x in Q20 bases; sum-of-contigs estimation

* NOTE: Estimated insert size may differ from sequence length
* (see http://www.hgsc.bcm.tmc.edu/docs/genbank_draft_data.html).
* NOTE: This is a 'working draft' sequence. It currently
* consists of 51 contigs. The true order of the pieces
* is not known and their order in this sequence record is
* arbitrary. Gaps between the contigs are represented as
* runs of N, but the exact sizes of the gaps are unknown.
* This record will be updated with the finished sequence
* as soon as it is available and the accession number will
* be preserved.
1 6683: contig of 6683 bp in length
6684 6783: gap of unknown length
6784 12614: contig of 5831 bp in length
12615 12714: gap of unknown length
12715 16628: contig of 3914 bp in length
16629 16728: gap of unknown length
16729 19332: contig of 2604 bp in length
19333 19433: gap of unknown length
19433 23463: contig of 4030 bp in length
23463 23562: gap of unknown length
23562 25954: contig of 2392 bp in length
25954 26054: gap of unknown length
26054 27990: contig of 1845 bp in length
27990 28000: gap of unknown length
28000 30376: gap of unknown length
30376 30476: contig of 2377 bp in length
30476 33824: gap of unknown length
33824 33924: contig of 3348 bp in length
33924 36144: gap of unknown length
36144 36245: contig of 2220 bp in length
36245 37544: gap of unknown length
37544 37644: gap of unknown length
37644 39820: contig of 1289 bp in length
39820 39920: gap of unknown length
39920 41519: contig of 2176 bp in length
41519 41619: gap of unknown length
41619 43294: contig of 1599 bp in length
43294 43394: gap of unknown length
43394 45184: contig of 1789 bp in length
45184 45284: gap of unknown length
45284 46781: contig of 1498 bp in length
46781 46882: gap of unknown length
46882 48928: contig of 1947 bp in length
48928 48929: gap of unknown length
48929 50954: gap of unknown length
50954 51054: contig of 2026 bp in length
51054 52624: gap of unknown length
52624 52724: contig of 1570 bp in length
52724 54323: gap of unknown length
54323 54423: contig of 1599 bp in length
54423 55973: gap of unknown length
55973 56073: contig of 1550 bp in length
56073 57987: gap of unknown length
57987 58087: contig of 1914 bp in length
58087 60090: gap of unknown length
60090 60190: contig of 2003 bp in length
60190 61729: gap of unknown length
61729 61829: contig of 1539 bp in length
61829 62932: gap of unknown length
62932 63032: contig of 1103 bp in length
63032 64638: gap of unknown length
64638 64738: contig of 1606 bp in length
64738 66604: gap of unknown length
66604 66704: contig of 1866 bp in length
66704 68585: gap of unknown length
68585 68685: contig of 1881 bp in length
68685 70000: gap of unknown length
70000 70100: contig of 1315 bp in length
70100 72122: gap of unknown length
72122 72222: contig of 2022 bp in length
72222 74384: gap of unknown length
74384 74484: contig of 2162 bp in length
74484 75864: gap of unknown length
75864 75864: contig of 1380 bp in length

* consists of 64 contigs. The true order of the pieces
 * is not known and their order in this sequence record is
 * arbitrary. Gaps between the contigs are represented as
 * runs of N, but the exact sizes of the gaps are unknown.
 * This record will be updated with the finished sequence
 * as soon as it is available and the accession number will
 * be preserved.

1
 7775: contig of 7775 bp in length
 7776: gap of unknown length
 7876: contig of 6367 bp in length
 14343: gap of unknown length
 14343: contig of 4925 bp in length
 19268: gap of unknown length
 19368: contig of 4397 bp in length
 23765: gap of unknown length
 23865: contig of 5346 bp in length
 29211: gap of unknown length
 29310: contig of 3865 bp in length
 33175: gap of unknown length
 33275: contig of 5325 bp in length
 33776: gap of unknown length
 38601: contig of 4201 bp in length
 42901: gap of unknown length
 43002: contig of 3183 bp in length
 46184: gap of unknown length
 46284: contig of 4398 bp in length
 50683: gap of unknown length
 50782: contig of 3223 bp in length
 54006: gap of unknown length
 54105: contig of 2981 bp in length
 57086: gap of unknown length
 57187: contig of 5085 bp in length
 62271: gap of unknown length
 62371: contig of 4323 bp in length
 66694: gap of unknown length
 66794: contig of 3528 bp in length
 70323: gap of unknown length
 70422: contig of 2713 bp in length
 73135: gap of unknown length
 73235: contig of 3878 bp in length
 77113: gap of unknown length
 77213: contig of 2386 bp in length
 79699: gap of unknown length
 82281: contig of 2592 bp in length
 82381: gap of unknown length
 84922: contig of 2531 bp in length
 85022: gap of unknown length
 87396: contig of 2374 bp in length
 87496: gap of unknown length
 89536: contig of 2040 bp in length
 89537: gap of unknown length
 89636: gap of unknown length
 92715: contig of 3079 bp in length
 92815: gap of unknown length
 95716: contig of 2901 bp in length
 95816: gap of unknown length
 97864: contig of 2048 bp in length
 97964: gap of unknown length
 100145: contig of 2181 bp in length
 100245: gap of unknown length
 103250: contig of 3005 bp in length
 103251: gap of unknown length
 103351: contig of 2301 bp in length
 105651: gap of unknown length
 105751: contig of 1560 bp in length
 107311: gap of unknown length
 107411: contig of 2477 bp in length
 109888: gap of unknown length
 112402: contig of 2414 bp in length
 112502: gap of unknown length
 113864: contig of 1362 bp in length
 113964: gap of unknown length
 116013: contig of 2049 bp in length
 116113: gap of unknown length

* 116114 118300: contig of 2187 bp in length
 * 118301 118400: gap of unknown length
 * 118401 120234: contig of 1834 bp in length
 * 120235 120334: gap of unknown length
 * 120335 122777: contig of 2443 bp in length
 * 122778 122877: gap of unknown length
 * 122878 125253: contig of 2376 bp in length
 * 125254 125353: gap of unknown length
 * 125354 127361: contig of 2008 bp in length
 * 127362 129536: contig of 2075 bp in length
 * 129537 129636: gap of unknown length
 * 129637 131660: contig of 2024 bp in length
 * 131661 131760: gap of unknown length
 * 131761 133298: contig of 1538 bp in length
 * 133299 133398: gap of unknown length
 * 133399 135392: contig of 1994 bp in length
 * 135393 135492: gap of unknown length
 * 135493 137378: contig of 1886 bp in length
 * 137379 137478: gap of unknown length
 * 137479 138815: contig of 1337 bp in length
 * 138816 138915: gap of unknown length
 * 138916 140791: contig of 1876 bp in length
 * 140792 140891: gap of unknown length
 * 140892 141919: contig of 1028 bp in length
 * 141920 142019: gap of unknown length
 * 142020 143557: contig of 1538 bp in length
 * 143558 143657: gap of unknown length
 * 143658 144951: contig of 1294 bp in length
 * 144952 145051: gap of unknown length
 * 145052 146516: contig of 1555 bp in length
 * 146517 146716: gap of unknown length
 * 146717 148415: contig of 1699 bp in length
 * 148416 148515: gap of unknown length
 * 148516 150775: contig of 2260 bp in length
 * 150776 150875: gap of unknown length
 * 150876 152549: contig of 1674 bp in length
 * 152550 152649: gap of unknown length
 * 152650 155221: contig of 2572 bp in length
 * 155222 155322: gap of unknown length
 * 155323 156794: contig of 1473 bp in length
 * 156795 156894: gap of unknown length
 * 156895 158288: contig of 1394 bp in length
 * 158289 158388: gap of unknown length
 * 158389 159553: contig of 1165 bp in length
 * 159554 159653: gap of unknown length

Query Match 4.4% Score 72.8; DB 2; Length 171322;
 Best Local Similarity 73.6%; Pred. No. 1e-10;
 Matches 106; Conservative 0; Mismatches 37; Indels 1; Gaps 1;

OY 182 GGGAGCTCCAGTATGAGACAGAAATGCTTACAGCTACATTCACAAACACACACACA 241
 DB 23241 GGGATCTTCACATGCGATGTAC-AGAAGTTGCTCACACACGCGACATACACACACA 23183
 OY 242 CACACACACACACACACACACACACACACACATGTTTCTTCCTCCCTCCACT 301
 DB 23182 CACACACACACACACACACACACACACACACACACTTCATTCCTTCGCCCTACTCG 23123
 OY 302 TCCCTCCCTTCCTTCGTGGTCCCA 325
 DB 23122 TCCACCCATCCCTGGGTTTCCTCA 23099

RESULT 12
 AC101368/c 72930 bp DNA linear HTG 23-NOV-2001
 LOCUS Mus musculus clone RP23-116F12, LOW-PASS SEQUENCE SAMPLING.
 AC101368
 AC101368.1 GI:17060143
 VERSION HTG: HTGS PHASEO.
 KEYWORDS house mouse.
 SOURCE ORGANISM Mus musculus

REFERENCE
AUTHORS
TITLE
JOURNAL
REFERENCE
AUTHORS

Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
1 (bases 1 to 72930)
Birren, B., Linton, L., Nusbaum, C. and Lander, E.
Mus musculus, clone RP23-116F12
Unpublished
2 (bases 1 to 72930)

Birren, B., Linton, L., Nusbaum, C., Lander, E., All, A., Allen, N.,
Anderson, S., Barina, N., Bastien, V., Boguslavsky, L., Boukhaalter, B.,
Brown, A., Camarata, J., Campopiano, A., Chang, J., Chazaro, B.,
Choepel, Y., Colangelo, M., Collins, S., Collamore, A., Cook, A.,
Cooke, P., Dearrellano, K., Dewar, K., Diaz, J.S., Dodge, S., Ferro, S.,
Ferreira, P., Fitzhugh, W., Gage, D., Galagan, J., Gardyna, S.,
Ginde, S., Gord, S., Goyette, M., Graham, L., Grand-Pierre, N.,
Hagos, B., Heaford, A., Horton, L., Hulme, W., Illey, I., Johnson, R.,
Jones, C., Kamat, A., Karatas, A., Kells, C., LaRoque, K.,
Lamazares, R., Landers, T., Lehoczy, J., Levine, R., Liu, G.,
Maclean, C., Macdonald, P., Major, J., Margulis, N., Matthews, C.,
McCarthy, M., McEwan, P., McKernan, K., McPeckers, R., Meldrum, J.,
Menues, L., Mihova, T., Mlenga, V., Murphy, T., Naylor, J., Nguyen, C.,
Norbu, C., Norman, C.H., O'Connor, T., O'Donnell, P., O'Neill, D.,
Oliver, J., Peterson, K., Phunkhang, P., Pierre, N., Pollara, V.,
Raymond, C., Retta, R., Rieback, M., Riley, R., Rise, C., Rogov, P.,
Roman, J., Rosetti, M., Roy, A., Santos, R., Schauer, S., Schupbach, R.,
Seaman, S., Severy, P., Spencer, B., Stange-Thomann, N., Stojanovic, N.,
Strauss, N., Subramanian, A., Talamas, J., Testaye, S., Theodore, J.,
Topham, K., Travers, M., Travis, N., Trigilio, J., Vassiliev, H.,
Viel, R., Vo, A., Wilson, B., Wu, X., Wyman, D., Ye, W.J., Young, G.,
Zainoun, J., Zembek, L., Zimmer, A. and Zody, M.

TITLE
JOURNAL
COMMENT

Direct Submission
Submitted (23-NOV-2001) Whitehead Institute/MIT Center for Genome
Research, 320 Charles Street, Cambridge, MA 02141, USA
Smit, A.F.A. & Green, P. (1996-1997)
http://ftp.genome.washington.edu/RM/RepeatMasker.html

Genome Center
Center: Whitehead Institute/ MIT Center for Genome Research
Center code: WIBR
Web site: http://www-seq.wi.mit.edu
Contact: sequence_submissions@genome.wi.mit.edu
Project Information
Center project name: L16561
Center clone name: 116_F_12

* NOTE: This record contains 89 individual
* sequencing reads that have not been assembled into
* contigs. Runs of N are used to separate the reads
* and the order in which they appear is completely
* arbitrary. Low-pass sequence sampling is useful for
* identifying clones that may be gene-rich and allows
* overlap relationships among clones to be deduced.
* However, it should not be assumed that this clone
* will be sequenced to completion. In the event that
* the record is updated, the accession number will
* be preserved.

1
729 828: contig of 728 bp in length
829 828: gap of 100 bp
1547 1646: contig of 718 bp in length
1647 2374: contig of 728 bp in length
2375 2474: gap of 100 bp
2475 3195: contig of 721 bp in length
3196 3295: gap of 100 bp
3296 4033: contig of 738 bp in length
4034 4133: gap of 100 bp
4134 4866: contig of 733 bp in length
4867 4966: gap of 100 bp
4967 5694: contig of 728 bp in length
5695 5794: gap of 100 bp
5795 6520: contig of 726 bp in length
6521 6620: gap of 100 bp
6621 7350: contig of 730 bp in length
7351 7450: gap of 100 bp

7451 8152: contig of 702 bp in length
8153 8252: gap of 100 bp
8253 8968: contig of 716 bp in length
8969 9068: gap of 100 bp
9069 9775: contig of 707 bp in length
9776 9875: gap of 100 bp
9876 10583: contig of 708 bp in length
10584 10683: gap of 100 bp
10684 11397: contig of 714 bp in length
11398 11497: gap of 100 bp
11498 12214: contig of 717 bp in length
12215 12314: gap of 100 bp
12315 13043: contig of 729 bp in length
13044 13143: gap of 100 bp
13144 13881: contig of 728 bp in length
13882 13981: gap of 100 bp
13982 14717: contig of 736 bp in length
14718 14817: gap of 100 bp
14818 15552: contig of 735 bp in length
15553 15652: gap of 100 bp
15653 16382: contig of 730 bp in length
16383 16482: gap of 100 bp
16483 17201: contig of 719 bp in length
17202 17301: gap of 100 bp
17302 18013: contig of 712 bp in length
18014 18113: gap of 100 bp
18114 18825: contig of 713 bp in length
18827 18926: gap of 100 bp
18927 19648: contig of 722 bp in length
19649 19748: gap of 100 bp
19749 20470: contig of 722 bp in length
20471 20570: gap of 100 bp
20571 21302: contig of 732 bp in length
21303 21402: gap of 100 bp
21403 22113: contig of 711 bp in length
22114 22213: gap of 100 bp
22214 22937: contig of 724 bp in length
22938 23037: gap of 100 bp
23038 23752: contig of 715 bp in length
23753 23852: gap of 100 bp
23853 24571: contig of 719 bp in length
24572 24671: gap of 100 bp
24672 25395: contig of 724 bp in length
25396 25495: gap of 100 bp
25496 26207: contig of 712 bp in length
26208 26307: gap of 100 bp
26308 27035: contig of 728 bp in length
27036 27135: gap of 100 bp
27136 27854: contig of 719 bp in length
27855 27954: gap of 100 bp
27955 28664: contig of 710 bp in length
28665 28764: gap of 100 bp
28765 29468: contig of 704 bp in length
29469 29568: gap of 100 bp
29569 30300: contig of 732 bp in length
30301 30400: gap of 100 bp
30401 31134: contig of 734 bp in length
31135 31234: gap of 100 bp
31235 31970: contig of 736 bp in length
31971 32070: gap of 100 bp
32071 32798: contig of 728 bp in length
32799 32898: gap of 100 bp
32899 33517: contig of 619 bp in length
33518 33617: gap of 100 bp
33618 34307: contig of 690 bp in length
34308 34407: gap of 100 bp
34408 35131: contig of 724 bp in length
35132 35231: gap of 100 bp
35232 35958: contig of 727 bp in length
35959 36058: gap of 100 bp
36059 36796: contig of 738 bp in length
36797 36896: gap of 100 bp
36897 37600: contig of 704 bp in length

37601	37700:	gap of	100 bp
37701	38421:	contig of 721 bp	in length
38422	38571:	gap of	100 bp
38572	39263:	contig of 742 bp	in length
39264	39363:	gap of	100 bp
39364	40085:	contig of 722 bp	in length
40086	40185:	gap of	100 bp
40186	40918:	contig of 733 bp	in length
40919	41018:	gap of	100 bp
41019	41750:	contig of 733 bp	in length
41751	41850:	gap of	100 bp
41851	42573:	contig of 723 bp	in length
42574	42673:	gap of	100 bp
42674	43404:	contig of 731 bp	in length
43405	43504:	gap of	100 bp
43505	44224:	contig of 720 bp	in length
44225	44324:	gap of	100 bp
44325	45038:	contig of 714 bp	in length
45039	45138:	gap of	100 bp
45139	45852:	contig of 714 bp	in length
45853	45952:	gap of	100 bp
45953	46678:	contig of 726 bp	in length
46679	46778:	gap of	100 bp
46779	47442:	contig of 714 bp	in length
47453	47552:	gap of	100 bp
47553	48295:	contig of 703 bp	in length
48296	48395:	gap of	100 bp
48396	49105:	contig of 710 bp	in length
49106	49205:	gap of	100 bp
49206	49927:	contig of 722 bp	in length
49928	50027:	gap of	100 bp
50028	50757:	contig of 730 bp	in length
50758	50857:	gap of	100 bp
50858	51575:	contig of 718 bp	in length
51576	51675:	gap of	100 bp
51676	52410:	contig of 735 bp	in length
52411	52510:	gap of	100 bp
52511	53225:	contig of 715 bp	in length
53225	53335:	gap of	100 bp
53326	54039:	contig of 714 bp	in length
54140	54139:	gap of	100 bp
54140	54850:	contig of 711 bp	in length
54851	54950:	gap of	100 bp
54951	55667:	contig of 717 bp	in length
55668	55767:	gap of	100 bp
55768	56478:	contig of 711 bp	in length
56479	56578:	gap of	100 bp

	Query Match	Best Local Similarity	4.3%;	Score 72.2;	DB 2;	Length 72930;
	Matches 107;	Conservative	0;	Mismatches 159;	Indels 0;	Gaps 0;
0y	138	AATATTGTTACAAAAGGTCGTTTCTCCCCACATGGACCCGGAGCTCCATGATA	197			
Db	47748	AATGAGATCTACGAGATCCAGTGAATCCCATCAAAATCCCAACCAATTTTTCACGATC	47689			
0y	198	TGGACAGATACCTTTACGTACATTCCTAAAACACACACACACACACACACACA	257			
Db	47688	AAGAAAGGTAATTTCAAACTCATATGGAACACACACACACACACACACACACA	47629			
0y	258	CACACACACACACACACACATGTTTCTCCCTCCCTCACTTCTCCCATTCCTGT	317			
Db	47628	CACACACACACACACACACAGAGAAATTCACCAANNNNNNNNNNNNNNNNNNNN	47566			
0y	318	GGTCCCAAGAGATGACCATATTGACTGTAGAATTCACACACCATAAAGCCCATCTGG	377			
Db	47568	NN	47509			
0y	378	GAGCATTTCCAGACTGATCTTTTGA	403			
Db	47508	NN	47483			

[illegible]

Elhaj,C., Escotto,M., Falls,T., Ferraguto,D., Flagg,N., Ford,J., Foster,P., Frantz,P., Gabisi,A., Gao,J., Garcia,A., Garner,T., Gatter,P., Gill,R., Gorrell,J.H., Guevara,W., Gunaratne,P., Hale,S., Hamilton,K., Harris,C., Harris,K., Hart,M., Haylak,P., Hanes,A., Hernandez,J., Hernandez,O., Hodgson,A., Hogues,M., Holloway,C., Hollins,B., Homsi,F., Howard,S., Huber,J., Huijck,S., Hume,J., Jackson,L.E., Jacobson,B., Jia,Y., Johnson,R., Jollivet,S., Joudan,S., Karlsson,E., Kelly,S., Khan,U., King,L., Kovach,J., Kovar,C., Kratovic,J., Kureshi,A., Landry,N., Leal,B., Lewis,L.C., Lewis,L., Li,J., Li,Z., Lichtharge,O., Lieu,C., Liu,J., Liu,W., Louised,H., Lozano,R.J., Lu,X., Lucier,A., Lucier,R., Luna,R., Ma,J., Maheshwari,M., Mapa,P., Martin,R., Martindale,A., Martinez,E., Massey,E., Mawhney,E., McLeod,M.P., Meador,M., Mel,G., Metzger,M., Miner,G., Miner,Z., Mitchell,T., Mohabhat,K., Morgan,M., Morris,S., Moser,M., Newton,J., Newton,N., Nguyen,A., Nguyen,N., Nguyen,N., Nickerson,E., Nwokwenwo,S., Ogutu,N., Okunolu,G., Orangunye,N., Oviedo,R., Pace,A., Payton,B., Peery,J., Perez,L., Peters,L., Pickens,R., Primus,E., Pu,L.L., Quiles,M., Ren,Y., Rives,M., Rojas,A., Rojudoan,I., Rolfe,M., Ruiz,S., Saverly,G., Scherer,S., Scott,G., Shen,H., Shooshari,N., Slsson,H., Sutton,A., Syatek,A., Tabor,P., Tamerisa,A., Stanley,H., Tang,H., Tansey,J., Taylor,C., Taylor,T., Telford,B., Thomas,N., Thomas,S., Usmani,K., Vasquez,L., Vera,V., Villalon,D., Vinson,R., Walli,R., Wang,S., Ward-Moore,S., Warren,R., Washington,C., Wellington,S., Williams,G., Williamson,A., Wleczky,R., Wooden,S., Worley,K., Wu,C., Wu,Y., Wu,Y.F., Zhou,J., Zorrilla,S., Nelson,D., Weinstein,G. and Gibbs,R.

Unpublished

2 (bases 1 to 176047)

Worley,K.C.

Direct Submission

Submitted (09-NOV-2001) Human Genome Sequencing Center, Department of Molecular and Human Genetics, Baylor College of Medicine, One Baylor Plaza, Houston, TX 77030, USA
On Dec 21, 2001 this sequence version replaced gi:1687456.

Genome Center

Center: Baylor College of Medicine

Center code: BCM

Web site: <http://www.hgsc.bcm.tmc.edu/>

Contact: hgsc-help@bcm.tmc.edu

Project Information

Center project name: G6NG

Center Clone name: CH230-189013

Summary Statistics

Assembly program: Phrap; version 0.990329First call to findPhrapList

Consensus quality: 151055 bases at least Q40

Consensus quality: 158732 bases at least Q30

Estimated insert size: 165129 bases at least Q20

Quality coverage: 0x in Q20 bases; sum-of-contigs estimation

Quality coverage: 2.5x in Q20 bases; sum-of-contigs estimation

NOTE: Estimated insert size may differ from sequence length

NOTE: <http://www.hgsc.bcm.tmc.edu/docs/Genbank.draft.data.html>.

NOTE: This is a 'working draft' sequence. It currently

consists of 42 contigs. The true order of the pieces

is not known and their order in this sequence record is

arbitrary. Gaps between the contigs are represented as

runs of N, but the exact sizes of the gaps are unknown.

This record will be updated with the finished sequence

as soon as it is available and the accession number will

be preserved.

1 12798: contig of 12798 bp in length

* 12799 12898: gap of unknown length

* 12899 25237: contig of 12339 bp in length

* 25238 25337: gap of unknown length

* 25338 37905: contig of 12578 bp in length

* 37906 38005: gap of unknown length

* 38006 45387: contig of 7382 bp in length

* 45388 45487: gap of unknown length

* 45487: gap of unknown length

45488 52837: contig of 7350 bp in length
* 52838 52937: gap of unknown length
* 52938 60684: contig of 7747 bp in length
* 60685 60785: gap of unknown length
* 60785 67091: contig of 6307 bp in length
* 67091 67191: gap of unknown length
* 67191 73887: contig of 6696 bp in length
* 73887 73988: gap of unknown length
* 73988 82832: contig of 8844 bp in length
* 82832 82932: gap of unknown length
* 82932 87168: contig of 4237 bp in length
* 87168 87269: gap of unknown length
* 87269 92059: contig of 4790 bp in length
* 92059 92159: gap of unknown length
* 92159 97934: contig of 5775 bp in length
* 97934 98033: gap of unknown length
* 98033 101431: contig of 3398 bp in length
* 101431 101531: gap of unknown length
* 101531 105159: contig of 3627 bp in length
* 105159 105259: gap of unknown length
* 105259 110382: contig of 5124 bp in length
* 110382 110482: gap of unknown length
* 110482 115931: contig of 5449 bp in length
* 115931 116031: gap of unknown length
* 116031 119788: contig of 3757 bp in length
* 119788 119888: gap of unknown length
* 119888 123331: contig of 3443 bp in length
* 123331 123431: gap of unknown length
* 123431 126327: contig of 2895 bp in length
* 126327 126427: gap of unknown length
* 126427 128477: contig of 2050 bp in length
* 128477 128577: gap of unknown length
* 128577 131745: contig of 3169 bp in length
* 131745 131845: gap of unknown length
* 131845 131846: contig of 4266 bp in length
* 131846 136211: gap of unknown length
* 136211 136212: contig of 3326 bp in length
* 136212 139537: gap of unknown length
* 139537 139637: contig of 2196 bp in length
* 139637 141833: gap of unknown length
* 141833 141933: gap of unknown length
* 141933 144814: contig of 2881 bp in length
* 144814 144914: gap of unknown length
* 144914 148363: contig of 3449 bp in length
* 148363 148463: gap of unknown length
* 148463 150147: contig of 1684 bp in length
* 150147 150247: gap of unknown length
* 150247 152674: contig of 2427 bp in length
* 152674 152775: gap of unknown length
* 152775 154591: contig of 1817 bp in length
* 154591 154691: gap of unknown length
* 154691 156159: contig of 1468 bp in length
* 156159 156259: gap of unknown length
* 156259 157478: contig of 1219 bp in length
* 157478 157578: gap of unknown length
* 157578 158858: contig of 1280 bp in length
* 158858 158959: gap of unknown length
* 158959 161046: contig of 2088 bp in length
* 161046 161146: gap of unknown length
* 161146 163452: contig of 2306 bp in length
* 163452 163552: gap of unknown length
* 163552 166098: contig of 2546 bp in length
* 166098 166198: gap of unknown length
* 166198 167902: contig of 1704 bp in length
* 167902 168002: gap of unknown length
* 168002 169409: gap of unknown length
* 169409 169509: contig of 1407 bp in length
* 169509 170627: contig of 1118 bp in length
* 170627 170727: gap of unknown length
* 170727 172694: contig of 1967 bp in length
* 172694 172794: gap of unknown length
* 172794 173803: contig of 1009 bp in length
* 173803 173903: gap of unknown length
* 173903 174905: contig of 1002 bp in length

TITLE
JOURNAL
REFERENCE
AUTHORS
TITLE
JOURNAL

COMMENT

